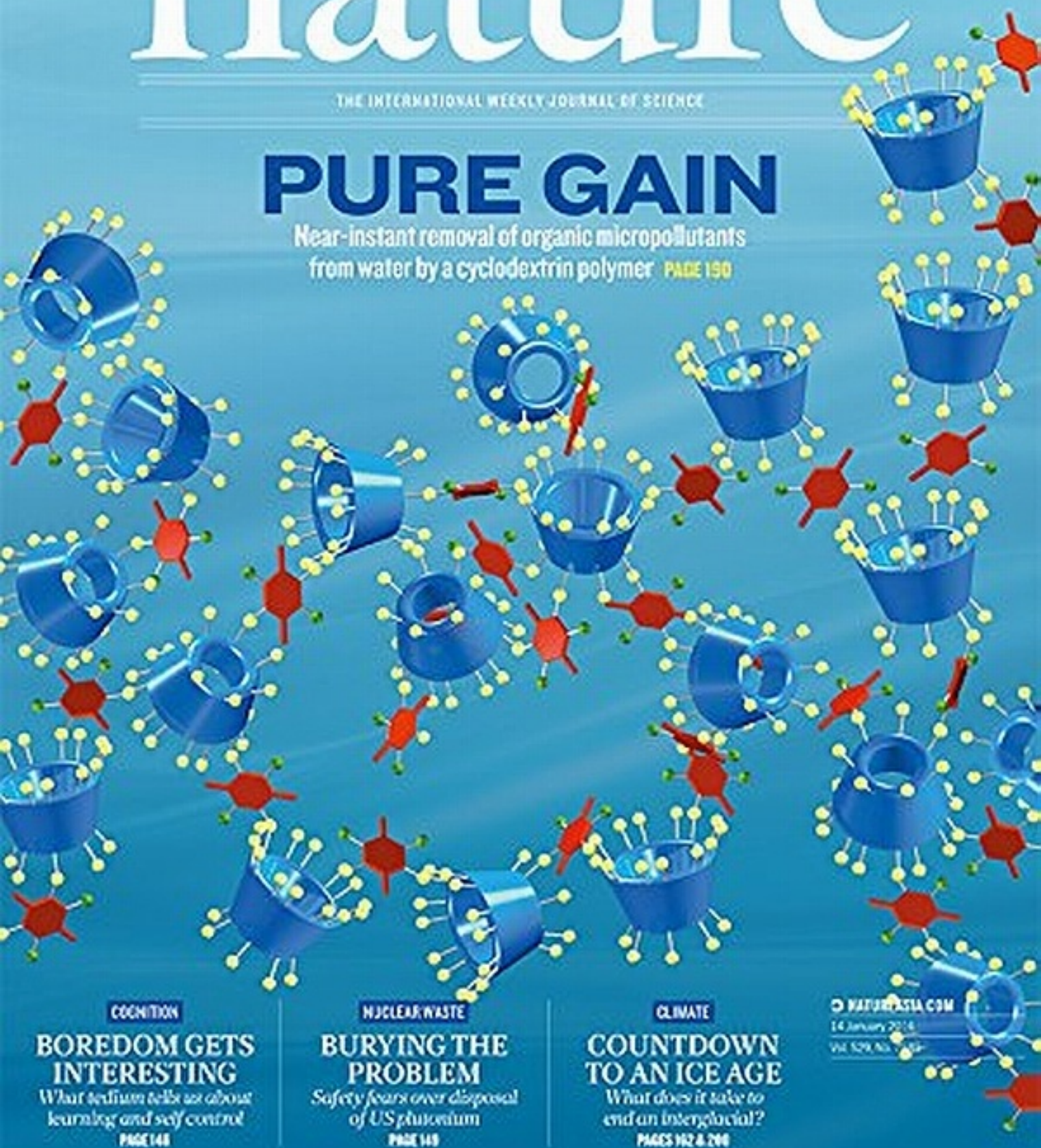


nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

PURE GAIN

Near-instant removal of organic micropollutants from water by a cyclodextrin polymer **PAGE 150**



COGNITION

BOREDOM GETS INTERESTING

What fMRI tells us about learning and self-control

PAGE 148

NUCLEAR WASTE

BURYING THE PROBLEM

Safety fears over disposal of US plutonium

PAGE 149

CLIMATE

COUNTDOWN TO AN ICE AGE

What does it take to end an interglacial?

PAGES 162 & 206

NATURE.COM

14 January 2016

Vol. 529, No. 7554

THIS WEEK

EDITORIALS

CHANGES Three new titles join the Nature journal club **p.128**

FAME New chemical elements should honour Primo Levi **p.129**

FANTASTIC VOYAGE Fossil fins show secrets of swimming squid **p.131**



Solving the drink problem

The United Kingdom's new guidelines on alcohol consumption are a sound example of evidence-based policymaking.

In his landmark song 'Heroes', David Bowie sang: "I, I'll drink all the time." Alcohol played such a part in Bowie's life that many tributes have taken care to point out that the musician was a non-drinker at the time of his death at the weekend.

Britain has a curious relationship with alcohol, as generations of visitors from abroad have experienced and pondered first-hand on any given evening. Whereas the people of other countries might drink to be sociable or as part of a meal, large numbers of Britons, many have observed, tend to drink alcohol like someone is trying to take it away.

Well, now somebody is — at least according to the reaction of some media commentators to last week's shift in official government guidelines on how much alcohol consumption is advisable. Just in time to reinforce any wavering new-year pledges to cut down on drinking, the UK Chief Medical Officers announced that neither men nor women should consume more than 14 units of alcohol a week — around 7 glasses of wine or 6 pints of average-strength beer. For British men, the amount is substantially less than the previous maximum guideline of 21 units per week. (The new advice is, at this stage, only draft guidance.) The guide amount is also less than comparable advice issued by many other nations.

Predictably, most dissent focused on the political argument that the government has no business telling people how to live their lives, and, presumably, speed their own deaths. Right-wing UK politician Nigel Farage led the (only just tongue-in-cheek) calls for those outraged by the latest example of "nanny state" politics to protest by heading immediately to the pub.

Disagreement with the scientific and medical basis for the new guidelines was more half-hearted. Most people in Britain seem to grudgingly accept that drinking too much is a bad thing, just as they have for a series of antisocial and unhealthy behaviours targeted in recent times — driving without seatbelts, supermarkets placing racks of chocolate at tills at child-friendly heights, and smoking, for instance. (This is a nation, remember, that felt it had to point out in official guidance as recently as 1984 that 56 drinks in a single week was "too much".)

In fact, despite some attempts to whip up outrage, there are signs that the British government is pushing against an open door in its attempts to get people to drink less. Alcohol consumption is reportedly falling, the number of people who abstain entirely is increasing, and the plague of young binge drinkers is in decline.

The statement that there is no 'safe' level of alcohol consumption is a solid one. Those who wish to dispute this should first read the evidence produced by the guidelines development group for the Chief Medical Officers, which includes modelling to balance risks and benefits (see go.nature.com/aauzdp). It shows that the past 20 years have produced a wealth of new evidence strongly linking alcohol use to cancer risk. And — contrary to the legion of newspaper stories — the minor health benefits of drinking are realized only by women over the age of 55, and then only at very low consumption levels. Red wine won't save you

from occasionally having to take a bit of exercise.

Decades hence, society may look back at today's acceptance and even celebration of alcohol and shake its collective head in the same way that we now view the acceptance of tobacco smoking, or the use of opium as a tonic.

Having an evidence-based recommendation is one thing. Actually changing behaviour is quite another. Millions of British men and women admit to routinely drinking more than they should. A sizeable fraction of those still drink more than 50 units a week. And the UK experts also pointed out the (not so) sobering fact that behavioural experts "found little evidence regarding the impact of any guidelines in changing health behaviours".

Still, it is a starting point, and the scientists whose work fed into the new guidelines should be proud. Converting solid evidence into scientifically grounded policy is something that everyone can raise a glass to. And more people now have the evidence to decide for themselves what type of drink should go into it. ■

"The scientists whose work fed into the new guidelines should be proud."

A secure future

Research advances mean that the time is ripe to ratify the ban on testing nuclear weapons.

This year marks the twentieth anniversary of the Comprehensive Nuclear-Test-Ban Treaty (CTBT) agreement, so the timing of the latest nuclear blast from North Korea is pertinent. The country's continued testing — this is its fourth test since 2006 — puts it on a path to developing miniaturized warheads that could be placed on missiles, risking an arms race in the region and increased global instability.

North Korea is one reason why the CTBT is not yet in force. The dictatorship is one of eight nuclear-capable nations that have yet to ratify the agreement, along with China, Egypt, India, Iran, Israel, Pakistan and the United States.

Science may seem to have little leverage in the volatile mixture of global power struggles and regional stand-offs, but it has been successful before. A major reason that so many countries were willing to sign up to the treaty in 1996 was the diligent research by a group of international scientists — known as the Group of Scientific Experts — established 20 years earlier in 1976. It had drawn up a credible road map of what technologies would be needed to verify that no country could cheat on its treaty obligations by carrying out undetected tests, thus giving them a military edge on those who abided by the rules.

Scientists can help again now — not least by explaining to politicians that the United States' principal technical objections to ratification have been overcome. In 1999, the US Senate rejected then-president Bill Clinton's push for ratification by a 51–48 vote, with opponents unconvinced that the technology was ripe either to detect cheaters, or to ensure the reliability and safety of the vast existing US stockpile of nuclear weapons without explosive testing.

Given the intensity of partisan politics in Washington DC today, hopes of any renewed effort by the United States to ratify the CTBT might seem fanciful. But at a symposium organized by the US Department of Energy in October 2015, US Secretary of State John Kerry called for just that, saying that the administration was determined to “reopen and re-energize the conversation about the treaty”.

Backing the case for ratification at the symposium were leading government scientists, such as US energy secretary Ernest Moniz — who had a key role in brokering the deal between the West and Iran over that country's nuclear programme last July — and the heads of US nuclear-weapons labs at the Lawrence Livermore, Los Alamos and Sandia National Laboratories.

Kerry and the scientists pointed out that advances in research meant that the Senate's concerns from 1999 are no longer relevant. The detection within minutes of last week's nuclear test by North Korea once again demonstrates that the International Monitoring System of the Vienna-based Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization is up to the job it

was designed to do. The US Stockpile Stewardship Program for its nuclear weapons, established in 1995, has also shown that advances in computer simulations and other technologies can assure the safety and reliability of its stockpile without any nuclear testing.

Although the CTBT has yet to enter into force, it has set an international standard. With the exception of North Korea, all countries have refrained from nuclear testing since 1998, when India and Pakistan each carried out two nuclear tests.

The United States has an opportunity to show leadership. By ratifying the CTBT, it would put huge pressure on China, India, Pakistan and other countries to do likewise. Iran, having scored a major diplomatic success with its nuclear deal with six world powers, is also in a strong position to support ratification. That would leave the signature of North Korea, probably the most recalcitrant non-signatory, for the CTBT to be able to enter into force. But as the Iran deal and the Paris climate negotiations show, diplomacy can prevail in the most difficult circumstances.

The CTBT alone will not solve all the complex issues of possession of nuclear weapons — in particular the disingenuous refusal of nuclear-weapons states to respect their commitment to the 1970 Nuclear Non-Proliferation Treaty to make serious efforts to disarm. But ratification of the CTBT would be a crowning achievement for science-based evidence and diplomacy in nuclear disarmament. Scientists played a key part in underpinning the nuclear deal with Iran; they now need to help to convince politicians that the CTBT is another deal in the best interests of international security. ■

ANNOUNCEMENT

Three new Nature journals

The traditional stamping grounds of *Nature* and the *Nature* journals have been the fundamental sciences — the physical, chemical, biological, Earth and environmental sciences. Three journals launched this week restate our editorial and publishing commitment to these territories. And one of them also delves into other disciplines, especially the social sciences, in tackling some of the ‘grand challenges’ facing society.

Nature Energy is the journal with the broadest scope. Like *Nature Climate Change* and *Nature Plants*, it includes social science and policy research: the first issue features papers on ‘Policy trade-offs between climate mitigation and clean cook-stove access in South Asia’ and ‘Impacts of a 32-billion-gallon bioenergy landscape on land and fossil fuel use in the US’. But the journal is also committed to the natural sciences — and indeed to any research that assists humankind in getting to grips with the challenges of energy generation, storage and distribution. In short, *Nature Energy* will attend to how science, technologies and people can deliver, and are affected by, any and all energy systems.

Like *Nature* and all other *Nature* research and reviews journals, *Nature Energy*'s choice of what to publish lies entirely in the hands of its in-house editors, who are supported by external peer reviewers. Everyone on the editorial team (which includes a social scientist) sits in the same office and is able to work closely together in assessing submissions. This is of particular value when dealing with multidisciplinary submissions — a challenge that the journal sees as one of its principal missions.

Materials research is a key component of the energy-research landscape. It also contributes fundamental insights into materials themselves and provides contexts in which materials can be applied. High time, our editors and publishers concluded, that a

Nature journal should survey progress across all these fronts: hence this week's launch of *Nature Reviews Materials*. Like the other two journals, it is an online-only subscription journal.

The launch issue includes reviews that outline the computational design of energy materials, the latest advances in photovoltaic devices, the surface properties of superhydrophobic and icephobic materials, the synthesis of carbon nanostructures and the design of pro-angiogenic materials, which are valuable in combating cardiovascular disease. It also focuses on sustainable materials, immunotherapy materials and the history of nanotechnology and the electronics industry. *Nature Reviews Materials* aims to cover the making, measuring, modelling and manufacturing of materials — looking at materials all the way from laboratory discovery to their use in functional devices. And in the coming months, the journal will analyse the impact that materials research can make in the field of medicine and on our environment, ensuring a healthier and more sustainable future.

The third journal is *Nature Microbiology*. As the most abundant living entities on our planet, microorganisms are fundamental to every facet of life on Earth. *Nature Microbiology* is interested in all aspects of microorganisms, be it their evolution, physiology and cell biology; their interactions with each other, a host or an environment; or their societal significance. The editors of *Nature Microbiology* are keen for the journal to be inclusive of all types of microorganism, whether bacterial, viral, archaeal or eukaryotic in nature. Accordingly, the launch issue features articles on a diverse array of microorganisms and topics, including the speciation of wild yeasts by hybridization, the global distribution of and disease burden caused by a bacterium and the identification of a virus that borrows its capsid coat from another virus.

Increasingly, researchers, their funders — both public and private — and their institutions recognize that great research needs to be pursued in both fundamental and societally useful domains. Such research needs to be inclusive, in disciplinary terms, and to aim for the highest standards of robustness. It is our hope that the *Nature* group of journals can support these ambitions, and notably so in the launches this week. ■

RICHARD HAUGHTON



New chemistry revives elementary question

The periodic table is a public symbol of chemistry. But as it grows larger, we must stress that science is not just about producing lists, says Philip Ball.

Rarely does chemistry enjoy the limelight as it has in past weeks. The announcement by the International Union of Pure and Applied Chemistry (IUPAC) that the seventh row of the periodic table has been filled through the discovery of four artificially created elements (numbers 113, 115, 117 and 118) has excited wide public discussion. What will these substances be called? What chemical properties do they have? How much further can the periodic table be extended?

This enthusiastic reception is surely a boon for chemistry. But rarely, also, does a feted scientific discovery have so few implications for the research agenda. IUPAC's announcement is not even of a discovery as such, but of the organization's assessment that the claims for the elements pass muster, and of its judgement on whose claims take precedence. The handful of laboratories worldwide that are equipped for the awe-inspiring task of making new elements did not require this seal of approval before pressing further into uncharted terrain.

Every new superheavy element raises interesting questions: whether there exists a region in which nuclear stability increases rather than diminishes with increasing mass, for example, and whether relativistic effects of the ultrafast movement of electrons distort the repeating patterns of properties in the table. There is plenty to celebrate and to study.

Whether nuclear science is chemistry at all has been in dispute ever since it began. Ernest Rutherford considered it a great joke that his 1908 Nobel prize for exploring radioactive decay was in chemistry — an attempt, some say, to win nuclear science back to chemistry after Marie and Pierre Curie's work on radioactivity won them a share in the 1903 physics prize.

The case for calling it chemistry was strong in the days when isolating and analysing radionuclides depended on the skilful and inventive application of separation methods to tiny quantities of material. The same might be asserted for studying the properties of the superheavies today. The experiments that refined and characterized a few atoms apiece of elements 104 (rutherfordium) to 108 (hassium) — each decaying within tens of seconds at most — are breathtaking examples of ultra-sensitive chemical analysis. But the methods used to make the elements in the first place, bombarding heavy nuclei with heavy ions by accelerating the ions to energies capable of penetrating the repulsive electrostatic barrier around the target nuclei, fall squarely within high-energy physics.

A deeper issue is what popular interest in the new elements implies about the status of the periodic table itself. Its systematization of elements has made it an icon for chemistry as a whole. Yet chemists rarely need to refer to it, and most of them work with just a handful of the more common elements.

It is fair to say that the periodic table holds

more interest and glamour for the public than it does for the working chemist. That's awkward: it would seem to open a rift between what many people think chemistry is about (study of the elements) and what most chemists do (make molecules and materials, and investigate their properties and interactions).

There is, however, nothing unique about this. Tabulation or enumeration of fundamentals also features in physics (the particles of the standard model) and biology (the genetic code, lists of genes and taxonomy). These classification schemes loom large in the popular consciousness, so that physics is deemed to be about finding new particles (after the Higgs boson come supersymmetric particles, particles of dark matter and so on) and biology becomes about identifying 'genes for' certain traits.

An enthusiasm for list-making is understandable. Not only does it seem to make complex ideas simpler, but it brings order to chaos, and may genuinely point — as the periodic table and the standard model do — to underlying symmetries and principles. We all like a good system. But the danger is that science then starts to look like a 'piling up of facts' — a tendency that seems, in the age of big data, to be colouring public perception and infecting research agendas.

The challenge for chemists, then, is to find a way to capitalize on the allure and coherence of the periodic table while avoiding the impression that it somehow tells the story of their research.

This focus adds weight to the question of how the new elements will be named. It seems a pity that the parochialism and nationalism, bordering sometimes on chauvinism, of the past (see germanium, francium, scandium, americium

and various permutations on the Swedish town of Ytterby) seems likely to persist. ('Japonium' is widely anticipated for element 113, because priority for its discovery was awarded to a team at the Japanese research institute RIKEN.) Why not take the opportunity to awaken the imagination, rather than plant a flag?

I would dearly love to see an element called levium, after the writer and chemist Primo Levi. His *The Periodic Table* (Einaudi, 1975) remains the best book ever written about chemistry, and it would please my sense of irony to see a superheavy element given a name that could be interpreted as a reference to lightness.

Yet this is not just about levity. Levi's account of his time in the Auschwitz concentration camp, 1947's *If This Is a Man*, is one of the century's most profound and humane works, testament to fact that science can be a liberating, universal force for salvation, while recognizing its potential to be abused in terrible ways. Levium would signify that the periodic table is for all of humanity. ■

Philip Ball is a science journalist in London.
e-mail: p.ball@btinternet.com

WHY NOT TAKE THE
OPPORTUNITY TO
AWAKEN THE
IMAGINATION
RATHER THAN
PLANT A FLAG?

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/vbssbl

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

GENE EDITING

CRISPR fixes muscle disease

Three teams of researchers have used CRISPR–Cas9 gene editing to treat mice that have the most common and severe form of muscular dystrophy.

Duchenne muscular dystrophy is a fatal disease caused by mutations that disable the gene encoding dystrophin, an important muscle protein. Teams led by Charles Gersbach of Duke University in Durham, North Carolina; Amy Wagers of Harvard University in Cambridge, Massachusetts; and Eric Olson of the University of Texas Southwestern Medical Center in Dallas used the CRISPR–Cas9 gene-editing technique to repair the dystrophin gene in mice that have such mutations.

The three teams used viruses to shuttle the components of the CRISPR–Cas9 system into the muscle cells of infant and adult mice. Treated mice made functional dystrophin and showed improvements in cardiac and skeletal muscle function.

Science <http://doi.org/bbnp> (2016); *Science* <http://doi.org/bbbs> (2016); *Science* <http://doi.org/bbbs> (2016)

MATERIALS

Self-folding origami master

Heat can bend a thin polymer film into different shapes inspired by origami.

Previous self-folding materials could either bend themselves into a shape and return to their original form, or permanently change shape. Tao Xie at Zhejiang University in Hangzhou, China, and his colleagues created a material

that could do both. At a relatively low temperature of around 80 °C, the polymer's molecular chains shift but chemical bonds in the network remain intact, which causes the material to temporarily fold into a predefined shape. At a higher temperature of around 130 °C, the bonds break and reform, inducing a permanent change in the material's molecular structure.

The same polymer could fold into multiple different shapes, which might eventually be useful in devices

that are deployed in the body or in space.

Sci. Adv. 2, e1501297 (2016)

ANIMAL BEHAVIOUR

Sharks have a nose for navigation

Sharks use their keen sense of smell for navigation as well as for feeding.

Andrew Nosal at the Scripps Institution of Oceanography in La Jolla, California, and his colleagues plugged the noses of wild leopard sharks (*Triakis*

water temperatures reduced electricity production at mid-latitudes, where most of the world's electricity is generated. Annual usable power capacity decreased by 7–12% for thermoelectric plants and by 1.2–3.6% for hydroelectric plants in the 2050s.

The authors suggest that boosting the efficiency of power plants, along with other adaptation measures, could reduce these impacts.

Nature Clim. Change <http://doi.org/bbsp> (2016)



ENERGY

Electricity at risk in a warmer world

Global warming's effects on water availability could hamper electricity production at power plants worldwide in the coming decades.

Michelle van Vliet of Wageningen University in the Netherlands and her colleagues modelled electricity production throughout the twenty-first century at more than 24,000 hydroelectric facilities and at about 1,400 water-cooled thermoelectric plants powered by natural gas, coal or nuclear energy. Decreased stream flow and warmer

semifasciata) with cotton balls soaked in petroleum jelly, tagged the animals with acoustic transmitters and released them 9 kilometres offshore. Over roughly four hours, sharks without nose plugs swam two-thirds of the way back to shore in relatively straight paths, whereas sharks with plugged noses took more tortuous paths, swimming only one-third of the way back.

The sharks could be detecting gradients of chemicals that are associated with coastal marine life, such

as dissolved amino acids, the authors say.

PLoS ONE 11, e0143758 (2016)

CANCER BIOLOGY

Gene promotes melanoma spread

Suppressing a regulatory gene in skin cancer could block the spread of cancer cells throughout the body.

The gene, *BMI1*, has been linked to the growth of certain tumours. To study its effect on tumour spread, or metastasis, Jacqueline Lees of the Massachusetts Institute of Technology in Cambridge and her colleagues looked at melanoma tumours in mice. Melanoma cells that expressed high levels of *BMI1* were more likely to spread to the lungs than were tumours that had normal *BMI1* levels. The gene also promoted resistance to drugs, and induced the expression of genes that have been linked to invasive melanoma and poor disease prognosis in humans.

The results suggest that *BMI1* could be a compelling drug target, the authors say. **Genes Dev.** 30, 18–33 (2015)

INFECTIOUS DISEASE

Poliovirus tweaked for safer vaccines

Poliovirus has been genetically modified so that it can be used in vaccines



BIOL. LEFT.

LIN YIQUANG/XINHUA/EYEVINE

without the risk of spreading the disease.

Inactivated polio vaccine is currently made (pictured) using highly infectious strains of the virus. To guard against accidental release, the World Health Organization in Geneva, Switzerland, has called on manufacturers to switch to weakened live strains called Sabin strains, despite their tendency to mutate into infectious forms.

A team led by Philip Minor at the UK National Institute for Biological Standards and Control in Pottery Bar created a genetically modified Sabin strain that, when inactivated, still elicited an immune response in rats. However, the virus did not mutate into active forms in cell lines and failed to infect macaques, so it would be unlikely to spread the disease among humans if it was accidentally released.

PLoS Pathog. 11, e1005316 (2015)

HUMAN EVOLUTION

Immunity boosted by archaic humans

Genes inherited from ancient hominins have improved the human immune system.

Homo sapiens interbred with Neanderthals and other ancient humans called Denisovans less than 100,000 years ago. Janet Kelso and her team at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, looked for Neanderthal and Denisovan genetic ancestry that has benefited humans by analysing the genomes of hundreds of people from around the world. They found a cluster of three Toll-like receptor (TLR) genes, which are involved in rapidly sensing and responding to infections as part of the innate immune response. Two Neanderthal versions of this cluster and one from Denisovans are common in different human populations. The archaic TLR genes are linked to reduced

SOCIAL SELECTION

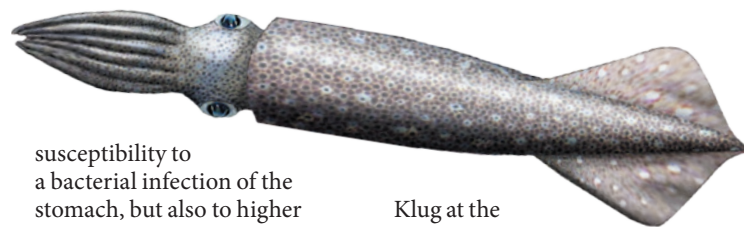
Popular topics on social media

Spoof kissing paper sparks debate

A satirical study showing that a mother's kisses didn't help injured children to feel better left several clues that it was fake. The funder was Proctor and Johnson, a made-up medical company, and one of the references was entitled, "So what the hell is going on here?". The paper, describing a fictional randomized controlled trial (RCT) of mothers kissing their toddlers, was designed to illustrate the limitations of evidence-based medicine, which uses data from such clinical trials to direct the practice of medicine. Many people who shared the article on Twitter played along with it. Angela Smith, a urologist at the University of North Carolina School of Medicine at Chapel Hill, tweeted: "Maternal kisses apparently ineffective at alleviating boo-boos in RCT—our household now switching to 'blowing on it.'" But some commentators said that the article, which the editor of the *Journal of Evaluation*

in *Clinical Practice* knowingly published in his journal, could be misleading and needs to be clearly labelled as satirical. **J. Eval. Clin. Pract.** <http://dx.doi.org/10.1111/jep.12508> (2015)

➔ **NATURE.COM**
For more on popular papers:
go.nature.com/e6rkaj



susceptibility to a bacterial infection of the stomach, but also to higher rates of allergies.

In a separate study, a team led by Lluís Quintana-Murci at the Pasteur Institute in Paris identified innate immunity genes that Europeans and Asians seem to have inherited from Neanderthals, including the same cluster of TLR genes. **Am. J. Hum. Genet.** <http://doi.org/bbn3> (2016); <http://doi.org/bbn2> (2016)

PALAEONTOLOGY

Squid relatives sped through water

Squid-like creatures that lived more than 60 million years ago could swim rapidly, supporting claims that they swam freely rather than just near the ocean bottom.

Fossils of belemnite marine animals from 200 million to 66 million years ago are common, but Christian

Klug at the University of Zurich in Switzerland and his colleagues report three *Acanthoteuthis* belemnite specimens from Germany with soft tissue components that have never been seen in such fossils before (reconstruction pictured). The tissue included fossil fins and organs called statocysts, which detect the direction and acceleration of movement through water. These suggest that the animals, which are relatives of modern squid, were fast-swimming predators.

This and other fossil evidence suggests a free-swimming, rather than an ocean-bottom-dwelling, lifestyle for belemnites. **Biol. Lett.** 12, 20150877 (2016)

➔ **NATURE.COM**
For the latest research published by Nature visit:
www.nature.com/latestresearch

SEVEN DAYS

The news in brief

PEOPLE

China science prize

A team led by quantum physicist Jian-Wei Pan was awarded the first-class prize of China's 2015 National Natural Science Award, one of the country's top science accolades, on 8 January. Pan and his team at the University of Science and Technology of China in Hefei won for their pioneering work in quantum entanglement and teleportation. For the first time in 11 years, no one was awarded China's top science prize, the State Supreme Science and Technology Award. Pharmacologist Youyou Tu, who last year won China its first science Nobel, had been tipped for the award.

BUSINESS

Pharma buyout

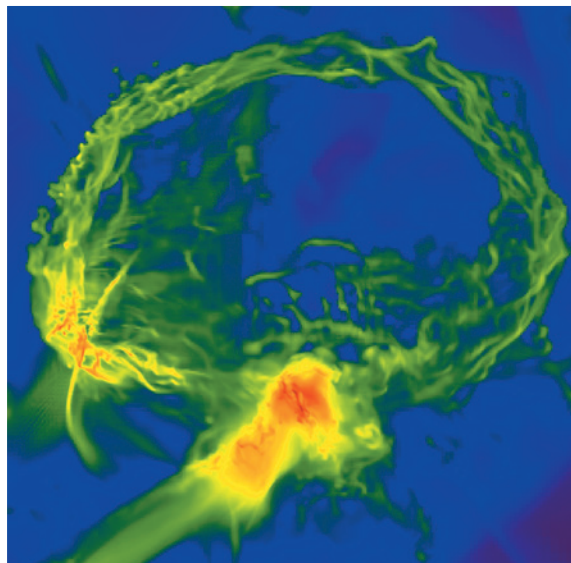
Pharmaceutical company Shire of Dublin is buying rival firm Baxalta of Bannockburn, Illinois, in a US\$32-billion deal, after a months-long pursuit. Both companies focus on rare-disease areas, including haematology, immunology and neuroscience. The firms say that as one company they will be able to make \$500 million in cost savings. Shire will

NUMBER CRUNCH

3.9×10^{13}

The number of bacteria in a typical human, alongside 3×10^{13} human cells. This new estimate challenges the idea that bacteria outnumber human cells by 10 to 1.

Source: Sender, R., Fuchs, S. & Milo, R. Preprint at bioRxiv <http://doi.org/bbpz> (2016).



Early star remnants

A faraway gas cloud has been discovered that contains tiny amounts of elements heavier than hydrogen and helium — such as carbon, oxygen and iron — that are possible remnants of the Universe's first stars. The elements were detected in spectra collected by the European Southern Observatory's Very Large Telescope in Chile, and computer simulations show how the Universe's first stars would have exploded and spewed the elements out (pictured). The results were reported at a meeting of the American Astronomical Society in Kissimmee, Florida, on 8 January. The cloud is so distant that it appears as it did 1.8 billion years after the Big Bang.

pay Baxalta shareholders in cash and shares, giving them around 34% ownership of the merged company. The deal is awaiting approval by regulators.

Cancer screening

The California sequencing-technology firm Illumina announced the formation of a new company, GRAIL, on 10 January. GRAIL will use Illumina's genetic-sequencing technology to screen for cancer from a blood sample. A 'liquid biopsy' would find minuscule amounts of tumour-specific DNA or RNA in the blood before the person

felt symptoms of the disease, when it may be easier to treat. GRAIL has more than US\$100 million in funding, in part from Bill Gates and from Amazon founder Jeff Bezos.

FUNDING

Singapore surge

Science spending in Singapore is set to surge by 18%, the government announced on 8 January. At its annual meeting, the country's Research, Innovation and Enterprise Council endorsed plans to invest 19 billion Singapore dollars (US\$13.2 billion)

between 2016 and 2020, up from 16.1 billion Singapore dollars between 2011 and 2015. The country will prioritize research funding in four areas: advanced manufacturing, health and biomedical sciences, services and the digital economy, and urban sustainability.

EVENTS

H-bomb claims

North Korea's fourth nuclear test on 6 January was almost certainly not a hydrogen bomb, contrary to the country's claims. The seismic event caused by the test was estimated at magnitude 4.85 by the Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization in Vienna. The explosion that caused that event was probably hundreds or thousands of times smaller than would have resulted from a hydrogen bomb, analysts say. North Korea might have tested a boosted fission device: a conventional fission bomb with a small quantity of the hydrogen isotopes tritium and deuterium added. See go.nature.com/gyqqya and page 127 for more.

Science passport

Seven science publishers, including PLOS and the American Geophysical Union, announced on 7 January that they will start requiring researchers to identify themselves using the ORCID (Open Researcher and Contributor ID) system when submitting papers. Globally, 1.8 million researchers have registered for ORCID's unique identifiers — machine-readable numbers akin to a scientific passport. The system is run by a non-profit organization that aims to create a transparent record

SMITH/WISE/O'SHEA/NORMAN/KHOCHFAR

linking scientists to their research outputs (see *Nature* **526**, 281–283; 2015).

Chimps returned

A legal battle over the ‘personhood’ of two chimpanzees has ended with their return to a primate facility in Louisiana, *Science* reported on 8 January. The two chimps were loaned to the State University of New York at Stony Brook for use as research animals. Animal-rights group the Nonhuman Rights Project sued in New York to have the animals released to a sanctuary, arguing that the chimps should have certain legal rights afforded to humans. The return of the chimps to the New Iberia Research Center in early December effectively removes the animals from New York’s jurisdiction.

Oil-pipeline fight

Pipeline firm TransCanada Corporation said on 6 January that it will seek more than US\$15 billion in compensation for economic losses under the North American Free Trade Agreement after the Keystone XL pipeline that it was due to build was cancelled (unused pipes **pictured**). The pipeline would have carried relatively dirty oil from tar sands in Alberta, Canada, to US refineries. But in November 2015, the US Department of State said that the project



was not in the “national interest”. TransCanada, which is headquartered in Calgary, called the decision “arbitrary and unjustified”, arguing that the project was environmentally benign. The company is also challenging the decision in the US federal court.

POLICY

Insecticide threat

The US Environmental Protection Agency (EPA) said on 6 January that the controversial insecticide imidacloprid does present a threat to bees and other pollinators. The preliminary risk assessment is the first of four on the neonicotinoids, an insecticide class that has been linked to bee declines. The European Food Safety Authority announced on 11 January that it would

be updating its own risk assessments of three neonicotinoids — clothianidin, thiamethoxam and imidacloprid. The European Union heavily restricted use of neonicotinoids in 2013 on the basis of previous evaluations.

UK drinking guides

Any level of alcohol intake increases cancer risk, according to draft guidelines released by the UK Chief Medical Officers on 8 January. Men and women should drink no more than 14 units of alcohol per week — around 7 glasses of wine or 6 pints of average-strength beer — according to the recommendations, which substantially lower the amount for men. The models used to calculate the recommendations considered risks and benefits, for instance cancer and alleged beneficial cardiovascular effects. The guidelines have

COMING UP

17 JANUARY

NASA plans to launch its Jason3 satellite to measure Earth’s sea levels, adding to knowledge of ocean circulation and climate change.

go.nature.com/rqfqmh

19–21 JANUARY

The Festival of Genomics takes place in London, bringing together industrialists, academics and policymakers.

go.nature.com/cw5hfb

18–22 JANUARY

PepTalk, dubbed ‘The Protein Science Week’, convenes in San Diego, California.

www.chi-peptalk.com

had a mixed reception, with some complaints that they are ‘nannying’. See page 127 for more.

FACILITIES

Linear collider

Japan should ramp up its expertise as it prepares to host the world’s next-generation particle smasher in the 2020s, reports the country’s High Energy Accelerator Research Organization (KEK) in Tsukuba. An action plan published on 6 January lays out the KEK’s goals for the preparation phases of the International Linear Collider, including a goal to triple the number of home-grown accelerator scientists and engineers. In 2012, Japanese researchers proposed hosting the 31-kilometre-long accelerator, which will smash electrons together with their antimatter partners. However, no government has yet promised any funding.

► **NATURE.COM**

For daily news updates see:

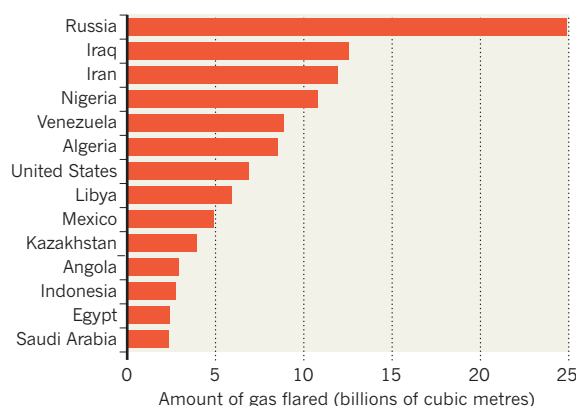
www.nature.com/news

TREND WATCH

Nations burned off around 143 billion cubic metres of natural gas — roughly 3.5% of global production — into the atmosphere in 2012, according to researchers at the US National Oceanic and Atmospheric Administration (C. D. Elvidge *et al.* *Energies* **9**, 14; 2016). Data from a polar-orbiting satellite showed that Russia led the way in terms of volume. The practice is common in fields that lack pipelines and markets for natural gas and policymakers are looking for ways to avoid the wastage.

3.5% OF NATURAL GAS WASTED IN FLARES

New satellite measurements track natural-gas flaring by country as policymakers seek to avoid wasting energy — and reduce emissions.



NEWS IN FOCUS

TAIWAN SARS epidemiologist poised to become vice-president **p.136**

EBOLA Scientists seek virus's origins as epidemic wanes **p.138**

PHYSICS Rumours frustrate India's bid for neutrino glory **p.140**



TIBET Bleak state of grasslands threatens millions of nomads **p.142**

KHAM/REUTERS/CORBIS



Vietnam's Viet-Laos cemetery contains the remains of thousands of people who died in the Vietnam War — but most are still unidentified.

FORENSIC SCIENCE

Vietnam begins huge effort to identify war dead

World's largest systematic identification project will use smart DNA-testing technology.

BY ALISON ABBOTT

Digging foundations for temples or schools, harvesting rice in paddy fields: these are some of the ways that the decaying remains of Vietnam War victims still turn up, 40 years after the conflict ended. Now an effort has begun that will use smart DNA technologies to identify the bones of the half a million or more Vietnamese soldiers and civilians who are thought still to be missing.

It is the largest ever systematic identification effort; only the identification of more than 20,000 victims of armed conflicts in Bosnia and

Herzegovina during the 1990s comes close.

"When I was a 21-year-old in the medical corps there, I never imagined that such a project could ever become possible," says Vietnam veteran and genomics pioneer Craig Venter, head of the J. Craig Venter Institute in La Jolla, California. "We thought of body counts as statistics — now, decades later, it may be possible to put names to them."

Although the United States has repatriated and identified most of its war dead, Vietnam has so far identified just a few hundred people, using outdated techniques. Yet people in Vietnam remain desperate to acquire the

remains of family members.

A few years ago, the government responded to their plight and asked the Advanced International Joint Stock Company (AIC) in Hanoi to investigate how best to proceed. The AIC consulted medical-diagnostics company Bioglobe in Hamburg, Germany, on how to equip the Vietnamese labs and train their scientists. In 2014, the Vietnamese government announced an investment of 500 billion dong (US\$25 million) in the project and said that it would upgrade its three existing DNA-testing centres.

This was great news, says Truong Nam Hai, head of the Institute of Biotechnology at ►

◀ the Vietnam Academy of Science and Technology, which hosts the first DNA-testing laboratory to be upgraded. In the 1990s, his institute proposed plans for identifying the missing, he says. However, “due to difficult circumstances at the time”, these did not take off.

Last month, the government signed a training and consultancy contract with Bioglobe, which will allow the sequencing effort to start.

“The technical challenges are considerable but tractable,” says Bioglobe’s chief executive, Wolfgang Höppner, who crafted the proposal for Vietnam. In the country’s hot and humid climate, DNA in bones that have lain in shallow graves for decades is likely to have degraded extensively. Moreover, contaminants from soil microbes can inhibit the enzymes that scientists use to amplify what little DNA remains to levels that can be analysed. And because of the large numbers of bones involved, the work needs to be done efficiently, adds Höppner.

Höppner’s proposal makes use of kits from Germany-based biotech company Qiagen, which have been designed to protect and reveal as much DNA as possible when dealing with difficult sources such as old, buried bones, and are also amenable to automated, ‘high throughput’ processes.

The identification process involves powdering bone samples and chemically breaking down their cells. Before amplification, the

DNA is extracted in sealed Qiagen cartridges that contain chemicals to wash away substances that could inhibit the process. Another Qiagen kit then checks the amplified DNA against a large set of genomic markers to create a DNA profile of the sample. The kit can also detect whether inhibitors are still present.

In cases in which inhibitors prove stubborn, samples will be analysed manually by slower, more complex methods that have been optimized by an experienced forensic laboratory run by the International Commission on Missing Persons (ICMP). That lab, in Bosnia and Herzegovina’s capital Sarajevo, led the effort to identify people killed during the 1990s conflict, including nearly all of the 8,000 or so who were massacred in 1995 in Srebrenica.

TRAINING BEGINS

The ICMP will also have a role in training Vietnamese scientists. Truong’s lab will next month send six scientists on a three-month programme. They will spend most of their time in Hamburg focusing on DNA tests, but they will also have a stint at the ICMP to learn other critical aspects of identification: how to avoid jumbling bones from different skeletons when exhuming them from mass graves, or how to look for clues in bones that might aid identification, such as pointers to height or gender.

It was possible to extract useful levels of DNA from around 80% of the bones from the

Srebrenica victims, says Thomas Parsons, head of the ICMP lab. The Vietnamese bones have been in the ground for longer and in a more damaging climate, but highly optimized methods and careful selection of skeletal samples will help, he says.

The Vietnam project will also need reference DNA from family members to compare with the bone DNA from victims. It plans to have an outreach programme calling for people to donate saliva samples to create a reference data bank — but this will not be easy. Many war victims may have died too young to have had children, and their parents may also be dead, so reference samples will have to come from more distant relatives whose DNA is less similar. “That is why it is particularly important to do the DNA analysis with a larger than normal set of markers,” says Höppner.

The outreach programme will also call for people to come forward with information on where bones might be buried. Unlike in Bosnia, where investigators could in some cases use satellite imagery to identify mass graves, the Vietnamese effort will rely on witness reports, as well as on common and military knowledge.

Once all three government DNA-testing centres are upgraded, probably by 2017, they will together be able to identify between 8,000 and 10,000 people a year, says Truong. He also anticipates that the DNA project will improve Vietnam’s scientific culture. ■

POLITICS

Taiwan’s SARS hero is poised for vice-presidency

Epidemiologist who spearheaded response to outbreak is popular with scientists — and others.

BY DAVID CYRANOSKI IN TAIPEI

A famous and influential epidemiologist, Chen Chien-Jen, is set to become Taiwan’s vice-president after elections on 16 January.

If he does, it is hoped that Chen — an epidemiologist looked upon as a hero for his role in subduing Taiwan’s outbreak of severe acute respiratory syndrome (SARS) in 2003 — will help to infuse the new government with an air of integrity and collaboration, maintain good relations with China and stimulate ideas for revitalizing the economy.

“He can negotiate with anyone, and is always trying to help,” says the National Taiwan University’s president, Yang Pan-Chyr. “You wouldn’t think such a person would

be a candidate for a politician.”

Chen announced in November that he would be the running mate for Tsai Ing-Wen, leader of the Democratic Progressive Party (DPP). If Tsai were to win, it would be only the second time in Taiwan’s history that the ruling Kuomintang (KMT) party has been dethroned.

Tsai is ahead in all the polls: she leads the KMT candidate by 30 percentage points, according to the non-profit Cross-Strait Policy Association, which carries out research on relations between Taiwan and the mainland — and the KMT’s own survey puts her lead at 8 percentage points.

Chen, too, is popular — the Cross-Strait Policy Association puts his ‘admiration’ rating at 54%, compared with 27% for his counterpart

in the KMT. This is probably a result of his celebrity status with regard to the SARS epidemic.

Panic over the viral infection, which initially emerged in mainland China but quickly spread across many parts of the world, was exacerbated in Taiwan because the United Nations recognizes China’s claim that Taiwan is part of China, and thus refuses to give it an independent seat at meetings of the World Health Organization. Excluded from international discussions and sample sharing, Taiwan’s outbreak spiralled out of control even as authorities elsewhere were getting a grip on the epidemic.

It was Chen, who was appointed health minister as the epidemic was escalating in Taiwan, who headed containment efforts. He



Presidential candidate Tsai Ing-Wen has picked epidemiologist Chen Chien-Jen to be her running mate in the upcoming Taiwan election.

SOURCE: WEB OF SCIENCE

bolstered attempts to isolate patients so as to prevent spread in hospitals, and boosted screening for fever. Even today, mentioning his name can elicit an enthusiastic thumbs-up. “Chen is great,” a taxi driver in Taiwan told *Nature* in early January. “With SARS, he was so fast.”

Chen is also popular in scientific circles, where he is known for other groundbreaking work. His research on the effects of arsenic exposure led health agencies around the world to lower the levels deemed acceptable (C.-J. Chen *et al.* *Br. J. Cancer* **66**, 888–892; 1992), and his assessment of the risk of liver cancer in people with chronic hepatitis led to new treatment guidelines (C.-J. Chen *et al.* *J. Am. Med. Assoc.* **295**, 65–73; 2006). An online petition supporting Chen’s candidacy has received more than 1,600 signatures — including those of prominent academics. “Within days, hundreds of names poured in,” says Ming-Liang Lee, the former Taiwanese health minister who started the petition.

Many researchers value Chen’s personality. “He has the capacity and appeal to pull people together,” says Ming-Chu Hsu, chief executive of TaiGen, one of Taiwan’s most successful biotechnology companies.

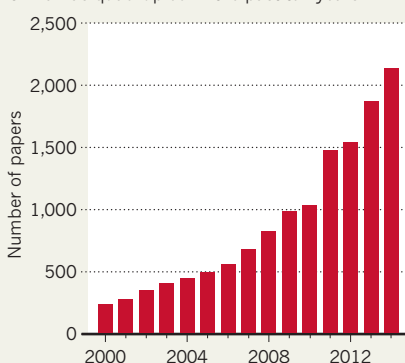
Chen carries a reassuring air of reliability. “He would be someone we can trust. Everyone seems to think so,” says Yang.

Chen himself told *Nature* that attributes honed during his time as a scientist — for example, the ability to solve problems — are beneficial to politics. He also said that it is crucial to revitalize Taiwan’s stagnant economy: increased competition in electronics from China and elsewhere has slashed the profits that once made Taiwan wealthy.

Tsai has outlined five areas in which Taiwan can innovate: biopharmaceuticals, green

CROSS-STRAIT COLLABORATION

The number of science papers co-authored by researchers from Taiwan and from mainland China has quadrupled in the past ten years.



energy, big data, precision machinery and national defence. To bolster those aims, Chen plans to establish a research system that encourages researchers and entrepreneurs to take risks. “Now the government doesn’t allow failure, so everyone goes for ‘me-too’ modifications, not innovation,” he told *Nature*.

Scientists and technology-based industrialists say that Chen and Tsai’s intention to promote innovation could bring a much-needed focus on Taiwanese science, although advocates are trying to keep things in perspective. “I think all science and technology would benefit from his taking office,” says Yang. “But maybe we are expecting too much.”

The DPP has traditionally emphasized Taiwanese autonomy, which riles Beijing, but “we don’t want to be troublemakers,” says Chen. He acknowledges that he himself came up against the Chinese authorities during the SARS epidemic, but says that agreement on

how to handle information on health and infectious diseases has largely resolved the issues.

A continuation of the status quo suits neuroscientist Chiang Ann-Shyn at Taiwan’s National Tsing Hua University; he expects Chen to act as an antidote to the DPP’s sometimes provocative statements on independence. “Relations with China have been good. I don’t think Chen will do anything radical,” he says. Two decades of stable relations following a crisis in the mid-1990s — when the mainland tested missiles in the strait — have led to a boom in business between Taiwan and the mainland, and research collaborations between them have quadrupled in the past ten years (see ‘Cross-strait collaboration’).

Hsu agrees with Chiang; her company’s antibiotic against multi-drug-resistant *Streptococcus pneumoniae* was the first drug developed in Taiwan to be submitted for approval on the mainland under new rules. “Health is one thing we can work on together,” she says.

Chen becomes emotional when talking about the possible end of his research career. Until recently, he had assumed that this would be at the nation’s premier research organization, the Academia Sinica, where he was vice-president until he declared himself Tsai’s running mate.

But although he was at first reluctant to join the electoral race, he finally decided that improving Taiwan’s social and economic situation was more important than his research.

A devout Catholic who consulted his archbishop before making his decision, Chen says that he considers his political career a “calling from God”. He adds: “I told the people in my laboratory that, for the coming years, it’s more important that I serve the people.” ■



Bats are suspected to be a wild reservoir of the Ebola virus.

EPIDEMIOLOGY

Ebola hunters go after viral hideout

As West Africa epidemic fades, researchers aim to prevent recurrences by finding the virus's natural host.

BY EWEN CALLAWAY

With the official end of Ebola transmission across West Africa anticipated on 14 January, an epidemic that killed more than 11,000 people in 2 years may be starting to fade into history. But that does not mean that Ebola has disappeared. The virus remains hidden in animal reservoirs, and is almost certain to spill over into humans again.

"We've got to focus on what could potentially happen next," says David Pigott, a spatial epidemiologist at the University of

Oxford, UK — and that means uncovering the species that harbour Ebola in the wild to try to prevent deadly outbreaks in the future.

It is no easy task. Since the disease first emerged in Zaire (now the Democratic Republic of the Congo) 40 years ago, efforts to trace the origins of the outbreaks, including the most recent one, have come up frustratingly empty. Wild gorillas and chimpanzees in central Africa have experienced occasional Ebola outbreaks. But like humans, these species are too ravaged by the virus to serve as its natural host. Experts say that a reservoir species is likely to harbour the virus

only at low levels, and without becoming sick.

The leading candidates are several species of fruit bat from across central and West Africa — where all known Ebola outbreaks have originated — that are often hunted for meat. A 2005 study¹ uncovered Ebola genetic material in some fruit bats from Gabon and the Democratic Republic of the Congo, and detected Ebola antibodies in the blood of others. Marburg virus, which is closely related to Ebola, is thought to be transmitted by fruit bats.

"I firmly believe fruit bats are the reservoir for Ebola," says Peter Daszak, a disease ecologist and president of EcoHealth Alliance, a conservation organization in New York City that plans to survey numerous bat species, including fruit bats, in Liberia for signs of Ebola infection.

Other researchers believe that focus is too narrow. "The evidence for fruit bats is the strongest, but it's still weak," says Fabian Leendertz, a wildlife epidemiologist at the Robert Koch Institute in Berlin.

Leendertz suspects another type of bat. He led a team that searched for the source of the latest West African outbreak in early 2014, a few months after a toddler in southern Guinea became the first human victim. The team captured dozens of bats near the toddler's village, but none — fruit-eating or otherwise — showed any conclusive signs of Ebola infection². Still, circumstantial evidence has led the researchers to suspect that the culprit may have been small insect-eating bats living in a tree near the toddler's home. Although the tree had burned down before researchers arrived, it had been filled with such bats, and villagers told the team that children often played in its hollowed-out trunk. The team is now looking more closely at insectivorous bats, but Leendertz cautions against focusing on any one animal.

UNUSUAL SUSPECTS

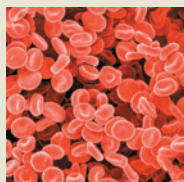
Some researchers advise casting the net even wider. "I don't buy the bat story for Ebola virus, not at all," says virologist Jens Kuhn of the US National Institute of Allergy and Infectious Diseases at Fort Detrick, Maryland. He thinks that bats are much too abundant and too closely associated with humans to explain an infection that has emerged just two dozen times over the past four decades. "It's going to be a strange host," he says. Even arthropods or fungi could be possibilities, he adds.

Others intend to look at more-familiar

PETE MULLER/NATL GEOGRAPHIC CREATIVE


**MORE
ONLINE**

TOP NEWS



Scientists bust myth that our bodies have more bacteria than human cells go.nature.com/2gxrbm

MORE NEWS

- The struggle to save the Middle East's cultural treasures go.nature.com/55vrco
- Famous 'Iceman' had familiar stomach infection go.nature.com/uncnwj
- Enzyme tweak boosts precision of CRISPR edits go.nature.com/3smg1p

NATURE PODCAST



Fibre and the microbiome, squeezing quantum states, and studying boredom nature.com/nature/podcast

SUSUMU NISHINAGA/SPL

species. The US Agency for International Development plans a two-year survey of animals ranging from rodents to livestock to domestic dogs and cats. These animals may not be natural reservoirs of Ebola, but they could contribute to spillovers into humans, says Dennis Carroll, director of the agency's Pandemic Influenza and Other Emerging Threats Unit.

But with so many question marks hovering over the identity of Ebola's reservoirs, some scientists say that it is time to eschew virus hunting in specific creatures and instead pursue more-holistic approaches that examine ecological and anthropological factors common to spillovers.

Tony Goldberg, an epidemiologist at the University of Wisconsin–Madison, is one such advocate. He no longer subscribes to the view that “we have to blanket the continent of Africa

with field-deployable DNA sequencers and sample everything that crawls, flies or swims and eventually we'll come across it. I used to think that way,” he says, “but I'm cooling off to that approach.”

His team is studying how bush-meat hunters interact with wild ecosystems to identify factors that might be linked to the spillover of zoonotic infections such as Ebola.

In a similar effort, a team led by Pigott and his colleague epidemiologist Simon Hay is looking at past outbreaks for common ecological factors, such as vegetation, elevation and the presence of suspected reservoir species such as fruit bats and carriers such as apes. By modelling these data, the team has created a map of areas at risk of Ebola spillovers³.

And Barbara Han, a disease ecologist at the

Cary Institute of Ecosystem Studies in Millbrook, New York, is using machine-learning techniques to predict which bat species are likely to harbour Ebola and related viruses because they share ecological factors common to suspected reservoir species.

Research on Ebola therapies and vaccines saw an infusion of public and private funding during the epidemic, and scientists hunting the virus in the wild hope to capture the same sense of urgency and financial support. But they know that the job won't be easy. “It has lit a fire under people's butts, mine included,” says Goldberg. “The problem is, we're not sure what to do with the fire.” ■

1. Leroy, E. M. *et al.* *Nature* **438**, 575–576 (2005).
2. Saéz, A. M. *et al.* *EMBO Mol. Med.* **7**, 17–23 (2014).
3. Pigott, D. M. *et al.* *eLife* **3**, e04395 (2014).

EXOPLANETS

Rebooted Kepler spacecraft hauls in the planets

Worlds found by K2 mission push beyond original discoveries.

BY ALEXANDRA WITZE, KISSIMMEE, FLORIDA

In the second phase of its life as a planet hunter, NASA's Kepler spacecraft is raking in exoplanet discoveries that are surprisingly different from those found during its first iteration.

Between 2009 and 2013, Kepler became the most successful planet-hunting machine ever, discovering at least 1,030 planets and more than 4,600 possible others in a single patch of sky. When a mechanical failure stripped the spacecraft of its ability to point precisely among the stars, engineers reinvented it in 2014 as the K2 mission, which looks at different parts of the cosmos for shorter periods of time.

In its first year of observing, K2 has netted more than 100 confirmed exoplanets, says astronomer Ian Crossfield at the University of Arizona in Tucson. They include a surprising number of systems in which more than one planet orbits the same star (E. Sinukoff *et al.* Preprint at <http://arxiv.org/abs/1511.09213>; 2015). The K2 planets are also orbiting hotter stars than are many of the Kepler discoveries.

“This is really showing the power and potential of K2,” says Crossfield. “These are things we never found with four years of Kepler data.” He and other scientists reported the findings last week at a meeting of the American Astronomical Society in Kissimmee, Florida.

The original Kepler mission was designed to answer a specific question: what fraction of Sun-like stars have Earth-sized planets around them? Unbound by those constraints — even if not as good at pointing itself — K2 has been able to explore wider questions of planetary origin and evolution. “Now we get to look at a much bigger variety,” says Steve Howell, the mission's project scientist at NASA's Ames Research Center in Moffett Field, California.

And because K2 looks at stars that are generally brighter and closer to Earth than Kepler did, the exoplanets that the mission finds are likely to be the best-studied for the foreseeable future. This is because they are near enough to allow astronomers to explore them with other telescopes on Earth and in space.

UNEXPECTED BOUNTY

In the past year, K2 has uncovered not just planets — such as three super-Earths orbiting a single star — but also surprises such as the disintegrating remains of a planet swirling around a white dwarf star. It has even probed exploding stars — because K2 stares constantly at a patch of the sky, it is able to catch a supernova as it brightens instead of later in its explosion, as other telescopes typically do.

Among the K2 planets confirmed so far, 58 are singletons, 28 come from systems with at least 2 planets and 14 are triples, Crossfield says. In addition, K2 has unearthed more than 200 candidate planets, says Andrew

Vanderburg, an astronomer at the Harvard Smithsonian Center for Astrophysics in Cambridge, Massachusetts.

K2 observes a larger fraction of the cool stars known as M dwarfs — the most common type of star in the Galaxy — than Kepler did. But surprisingly, fewer of the K2 planets are orbiting M-dwarf stars. A higher percentage of them, at least so far, circle stars that are hotter and more like the Sun, says Courtney Dressing, an astronomer at the California Institute of Technology (Caltech) in Pasadena.

K2 will begin a new type of planet-hunting on 7 April. Normally the spacecraft searches for a temporary dimming of a star caused when a planet crosses in front of it. For just under three months, however, it will look for the temporary brightening of cosmic objects, such as a galaxy, caused when a planet bends light as it crosses the line of sight between the object and the observer. The team expects to catch between 85 and 120 of these ‘microlensing’ planets during the campaign.

The survey will involve other telescopes and be the first automated search to be done simultaneously from the ground and in space, says Calen Henderson, an astronomer at NASA's Jet Propulsion Laboratory in Pasadena, California.

That means much more work ahead for mission scientists. “Kepler was one field and it ruined your summer,” says Caltech astronomer David Ciardi. “K2 is ruining our whole year.” ■



Physicists planning to build a neutrino detector in southern India have run into local opposition.

PHYSICS

Unfounded rumours delay Indian neutrino detector

Fears frustrate physicists in a global competition to understand elusive particles.

BY ELIZABETH GIBNEY

Time is running out for Indian scientists to build a facility that would let them compete in one of the hottest races in physics. The India-based Neutrino Observatory (INO) — an effort to learn about the masses and other properties of mysterious particles called neutrinos — is under threat as a result of baseless rumours about its aims and environmental impact. Despite a government go-ahead in January 2015 to build a massive detector under a mountain in the southern state of Tamil Nadu, opposition from environmentalists and state politicians means that not a single grain of earth has been shifted.

Neutrinos are abundant subatomic particles that are extremely hard to detect. Billions pass through each square centimetre of Earth every second, but barely any leave a trace. The INO would study neutrinos produced when cosmic rays strike the atmosphere, and would seek to reveal the relative masses of the three known types of neutrino. The measurements could lead to Nobel-prize-worthy insights into the relationship between nature's four fundamental

forces, as well as the imbalance between matter and antimatter in the Universe.

But if the INO is not built soon, other projects — including one that broke ground in China a year ago — may get there first, says D. Indumathi, a theorist at the Institute of Mathematical Sciences in Chennai who is part of the INO collaboration, and coordinates outreach for it. “Longer than a year of delay and I think it will be difficult to have viable physics goals, at least of the current type,” she says.

“They don’t know what the truth is, and I can understand that.”

Conceived in 2001 and originally slated for completion in 2012, the INO has faced a rocky path to construction. To shield the enormous detector from the confounding zoo of subatomic particles that pummels Earth’s surface, the facility needs to be built more than a kilometre underground. The first earmarked site was ruled out in 2009 after a lengthy battle with conservationists over its proximity to an elephant and tiger reserve.

The current site, in the Tamil Nadu district of Theni, faced opposition as soon as it was put

forward in 2010. Local villagers worried that the facility would deplete or contaminate their restricted water supply, and cut off access to land for grazing livestock, says Indumathi. But, she says, villagers consented after scientists assured them that the facility would not interfere with their resources.

Since then, however, local environmental organizations and regional politicians have taken up the issue, and the list of objections has swelled to include fears that the lab will emit radiation and store nuclear weapons, and that the excavation will threaten a nearby dam.

The rumours are untrue, says Naba Mondal, a physicist at the Tata Institute of Fundamental Research in Mumbai who leads the INO collaboration. INO scientists have visited schools and held community meetings to counter misconceptions. But many villagers have turned against the project. “They don’t know what the truth is, and I can understand that,” says Mondal.

At the root of the rumours is mistrust of the state and the scientific establishment, says Govind Krishnan, an Indian journalist who has closely followed the project. He believes that the fears that have been raised lie “in the

realm of fantasy”, but are understandable given the poor environmental record of past state-sponsored construction projects. Govind disagrees with activists who say that INO scientists have ignored the project’s impact on the poor, but he says that scientists’ efforts have been hampered by class and linguistic barriers.

India’s government allocated 15 billion rupees (US\$225 million) to construction when it gave the INO the green light last year, but the Madras High Court in Chennai brought the project to a standstill in March following a petition from local activists and politicians. The court said that the Tamil Nadu Pollution Control Board must give consent before construction can start. This is normally a routine, 45-day step, but the process has so far taken 9 months, says Mondal.

The politically contentious nature of the project means that the local board may well delay until after state elections in May. “I am confident that it will eventually be approved, but the question is when,” says Mondal. The delay is damaging the morale of students and researchers on the project, he adds.

Meanwhile, China expects to complete the Jiangmen Underground Neutrino Observatory in 2019. To remain competitive, the INO must start construction in the next few months, says Mondal. “Science is something you have to do in time. If you are not in time,

your results may not be that important.”

But neutrino physicists say that even if the INO loses the race, its findings would help to corroborate discoveries at other detectors. The INO takes a unique approach — using 50,000 tonnes of magnetized iron to separate atmospheric neutrino observations from their antineutrino counterparts. That will make its results interesting whenever they come out, says Mark Messier, a physicist at Indiana University Bloomington and co-spokesperson for the NOvA Neutrino Experiment at Fermilab

in Batavia, Illinois, which also has a chance of solving the neutrino-mass mystery.

Researchers point to other benefits, too. Putting a physics laboratory deep underground gives India the opportunity to host research into areas such as dark matter, they say — and it is empowering for Indian scientists to bring a major physics facility to fruition. “Already I’ve seen the tremendous difference it’s made to students having an experiment on which they call the shots,” says Indumathi. “So I really don’t care whether we get a Nobel prize or not.” ■

CORRECTIONS

The Editorial ‘Fishy limits’ (*Nature* **528**, 435; 2015) wrongly implied that the European Commission had set the fishing quotas. They were set by the Council of Ministers. The News story ‘Feuding physicists turn to philosophy’ (*Nature* **528**, 446–447; 2015) gave the wrong affiliation for Sabine Hossenfelder; she is now at the Frankfurt Institute for Advanced Studies. The News Feature ‘How to make the most of carbon dioxide’ (*Nature* **526**, 628–630; 2015) said that Carbon Recycling International produces 1.5% of global methanol; in fact, it makes 0.005%. The News Feature ‘Space. Time. Entanglement.’ (*Nature* **527**, 290–293; 2015)

wrongly said that Leonard Susskind began to think about computational complexity ten years ago — his work in the area began around three years ago. The News Feature ‘The truth about fetal tissue research’ (*Nature* **528**, 178–181; 2015) incorrectly stated that around 5.8 billion people have received vaccines made with the WI-38 and MRC-5 cell lines. In fact, companies have shipped some 5.8 billion vaccines made with these two cell lines. And a printing error meant that an earlier version of the News article ‘What to look out for in 2016’ (*Nature* **529**, 14–15; 2016) appeared that did not account for the fact that NASA has cancelled the 2016 launch of the Mars InSight probe.



A group of young Tibetan monks huddles on a degraded pasture on the Tibetan Plateau.

TROUBLE IN TIBET

Rapid changes in Tibetan grasslands are threatening Asia's main water supply and the livelihood of nomads.

BY JANE QIU

In the northern reaches of the Tibetan Plateau, dozens of yaks graze on grasslands that look like a threadbare carpet. The pasture has been munched down to bare soil in places, and deep cracks run across the snow-dusted landscape. The animals' owner, a herder named Dodra, emerges from his home wearing a black robe, a cowboy hat and a gentle smile tinged with worry.

"The pastures are in a bad state and lack the kind of plants that make livestock strong and grow fat," says Dodra. "The yaks are skinny and produce little milk."

His family of eight relies on the yaks for most of its livelihood

— milk, butter, meat and fuel. Dodra was forced to give up half of his animals a decade ago, when the Chinese government imposed strict limits on livestock numbers. Although his family receives financial compensation, nobody knows how long it will last.

"We barely survive these days," he says. "It's a hand-to-mouth existence." If the grasslands continue to deteriorate, he says, "we will lose our only lifeline".

The challenges that face Dodra and other Tibetan herders are at odds with glowing reports from Chinese state media about the health of Tibetan grasslands — an area of 1.5 million square kilometres — and the experiences of the millions of nomads there. Since the 1990s, the government has carried out a series of policies that moved once-mobile

herders into settlements and sharply limited livestock grazing. According to the official account, these policies have helped to restore the grasslands and to improve standards of living for the nomads.

But many researchers argue that available evidence shows the opposite: that the policies are harming the environment and the herders. “Tibetan grasslands are far from safe,” says Wang Shiping, an ecologist at the Chinese Academy of Sciences’ (CAS) Institute of Tibetan Plateau Research (ITPR) in Beijing. “A big part of the problem is that the policies are not guided by science, and fail to take account of climate change and regional variations.”

The implications of that argument stretch far beyond the Tibetan Plateau, which spans 2.5 million square kilometres — an area bigger than Greenland — and is mostly controlled by China. The grasslands, which make up nearly two-thirds of the plateau, store water that feeds into Asia’s largest rivers. Those same pastures also serve as a gigantic reservoir of carbon, some of which could escape into the atmosphere if current trends continue. Degradation of the grasslands “will exacerbate global warming, threaten water resources for over 1.4 billion people and affect Asian monsoons”, says David Molden, director general of the International Centre for Integrated Mountain Development (ICIMOD) in Kathmandu, Nepal.

Such concerns propelled me to make a 4,700-kilometre journey last year from Xining, on the northeastern fringe of the plateau, to Lhasa in the Tibetan heartland (see ‘Trek across Tibet’). Meeting with herders and scientists along the way, I traversed diverse landscapes and traced the Yellow and Yangtze Rivers to their sources. The trip revealed that Tibetan grasslands are far less healthy than official government reports suggest, and scientists are struggling to understand how and why the pastures are changing.

FENCED IN

It began to drizzle soon after we set off from the city of Xining on a stretch of newly built highway along the Yellow River. As our Land Cruiser climbed onto a 3,800-metre-high part of the plateau, the vista opened to reveal rolling hills blanketed by a thick layer of alpine meadow, resembling a gigantic golf course. We passed herds of sheep and yaks, white tents and nomads in colourful robes — along with barbed-wired fences that cut the rangeland into small blocks.

This part of the Tibetan Plateau, in a region known as Henan county, is blessed with abundant monsoonal rains every summer. The herders who live here are able to maintain healthy livestock and can make a decent living. “We have plenty to go around, and the livestock are well taken care of,” says herder Gongbu Dondrup.

But life has been different since the government began to fence up grasslands around a decade ago, says Dondrup. Before that, he took his herd to the best pastures at high elevations in the summer, and then came back down in the winter. Now, he must keep the yaks in an 80-hectare plot that the government assigned to his family. The pasture looks worn, and he is being pressed by the government to further downsize his herd. “I don’t know how long it can keep us going,” he says.

The fencing initiative is the latest of a string of Chinese grassland policies. After annexing Tibet in 1950, the young revolutionary Chinese republic turned all livestock and land into state properties. Large state farms competed with each other to maximize production, and livestock numbers on the plateau doubled over two decades, reaching nearly 100 million by the late 1970s. But in the 1980s, as China moved towards a market-based economy, Beijing swung to the other extreme: it privatized the pastures and gave yaks back to individual households, hoping that the move would push Tibetans to better manage their land and so boost its productivity.

Despite the privatization, nomads continued to use the rangeland communally — often in groups led by village elders. Then the government began to limit herds, and it built fences to separate households and villages. “This has totally changed the way livestock are traditionally raised on the plateau, turning a mobile lifestyle into a sedentary existence,” says Yang Xiaosheng, director of Henan county’s

rangeland-management office.

The fencing policy does have merits when applied in moderation, says Yönten Nyima, a Tibetan policy researcher at Sichuan University in Chengdu. Because an increasing number of nomads now lead a settled life — at least for parts of the year — it helps to control the level of grazing in heavily populated areas, he says. “Fencing is an effective way to keep animals out of a patch of meadow.” Many herders also say that it makes life much easier: they do not have to spend all day walking the hills to herd their yaks and sheep, and if they go away for a few days, they don’t worry about the animals running off.

But the convenience comes at a cost, says Cao Jianjun, an ecologist at Northwest Normal University in Lanzhou. Fenced pastures often show signs of wear after a few years. In a 2013 study, Cao and his colleagues measured growth of the sedge species preferred by livestock in two scenarios: enclosed pastures and much larger patches of land jointly managed by up to 30 households. Despite similar livestock densities in both cases, the sedge grew twice as fast in the larger pastures, where animals could roam and plants had more opportunity to recover¹. That matches the experience of Henan county herders, who say that their land sustains fewer animals than it has in the past.

WATER WORRIES

The future of the grasslands looked even bleaker as we left relatively well-to-do Henan county and ventured into the much higher, arid territory to the west. After 700 kilometres, we reached Madoi county, also known as *qianhu xian* (‘county of a thousand lakes’), where the Yellow River begins. Although this region gets only 328 millimetres of rain on average each year, about half of what Henan receives, Madoi was once one of the richest counties on the plateau — famous for its fish, high-quality livestock and gold mines.

Now, the wetlands are drying up and sand dunes are replacing the prairies, which means that less water flows into the Yellow River. Such changes on the plateau have contributed to recurring water shortages downstream: the Yellow River often dries up well before it reaches the sea, an event not recorded before 1970.

In 2000, China sought to protect this region, along with adjacent areas that give rise to the Yangtze and Mekong Rivers, by establishing the Sanjiangyuan (or Three-Rivers’ Headwaters) National Nature Reserve, an area nearly two-thirds the size of the United Kingdom.

“THE POLICIES ARE NOT GUIDED BY SCIENCE, AND FAIL TO TAKE ACCOUNT OF CLIMATE CHANGE AND REGIONAL VARIATIONS.”

Nearly one-tenth of the reserve area falls into core zones in which all activities, including herding, are prohibited. The government spends hundreds of millions of US dollars each year on moving nomads out of those core areas, constructing steel meshes to stabilize the slopes and planting artificially bred grass species to restore the eroded land. Outside the core regions, officials have banned grazing on ‘severely degraded grasslands’, where vegetation typically covers less than 25% of the ground. Land that is ‘moderately degraded’, where vegetation coverage measures 25–50%, can be grazed for half of the year.

Such policies — and related initiatives to limit livestock numbers and fence off areas of pasture — have not been easy on the herders, says Guo Hongbao, director of the livestock-husbandry bureau in Nagchu county in the southern Tibetan Plateau. “The nomads have made sacrifices for protecting the grasslands,” he says. But he also says that the strategies have paid off. Guo and other officials point to satellite studies showing that the plateau has grown greener in the past three decades². This increase in vegetation growth, possibly the result

of a combination of grazing restrictions and climate change, “has had a surprisingly beneficial effect on climate by dampening surface warming”, says Piao Shilong, a climate modeller at Peking University.

But ecologists say that such measurements look only at surface biomass and thus are not a good indicator of grassland health. “Not all vegetation species are equal,” says Wang. “And satellites can’t see what’s going on underground.”

This is particularly important in the case of the sedge species that dominate much of the Tibetan Plateau, and that are the preferred food of livestock. These species, part of the *Kobresia* genus, grow only 2 centimetres above the surface and have a dense, extensive root mat that contains 80% of the total biomass.

Studies of pollen in lake sediments show that *Kobresia* and other dominant sedges emerged about 8,000 years ago, when early Tibetans began burning forests to convert them to grasslands for livestock³. The prehistoric grazing helped to create the thick root mat that blankets the vast plateau and that has stored 18.1 billion tonnes of organic carbon.

But *Kobresia* plants are being driven out by other types of vegetation, and there is a risk that the locked-up carbon could be released and contribute to global warming. Every now and then on the trip to Lhasa, we passed fields blooming with the beautiful red and white flowers of *Stellera chamaejasme*, also known as wolf poison. “It’s one of a dozen poisonous species that have increasingly plagued China’s grasslands,” says Zhao Baoyu, an ecologist at the Northwest Agriculture and Forestry University in Yangling. Zhao and his colleagues estimated that poisonous weeds have infested more than 160,000 square kilometres of the Tibetan grasslands, killing tens of thousands of animals a year⁴.

Herders also report seeing new grass species and weeds emerge in recent years. Although most are not toxic, they are much less nutritious than *Kobresia* pastures, says Karma Phuntsho, a specialist on natural-resource management at ICIMOD. “Some parts of the plateau may seem lush to an untrained eye,” he says. “But it’s a kind of ‘green desertification’ that has little value.”

In one unpublished study of the northeastern Tibetan Plateau, researchers found that *Kobresia* pastures that had gone ungrazed for more than a decade had been taken over by toxic weeds and much taller, non-palatable grasses: the abundance of the sedge species had dropped from 40% to as low as 1%. “*Kobresia* simply doesn’t stand a chance when ungrazed,” says Elke Seeber, a PhD student at the Senckenberg Natural History Museum in Görlitz, Germany, who conducted the field experiment for a project supported by the German Research Foundation (DFG).

“HAVING A SWEEPING GRAZING POLICY REGARDLESS OF GEOGRAPHICAL VARIATIONS IS A RECIPE FOR DISASTERS.”

The changes in vegetation composition have important implications for long-term carbon storage, says project member Georg Guggenberger, a soil scientist at Leibniz University of Hanover in Germany. In moderately grazed *Kobresia* pastures, up to 60% of the carbon that is fixed by photosynthesis went into the roots and soil instead of the above-ground vegetation — three times the amount seen in ungrazed plots⁵. This underground organic carbon is much more stable than surface biomass, which normally decomposes within a couple of years and releases its stored carbon into the air. So a shift from *Kobresia* sedge to taller grasses on the plateau will ultimately release a carbon sink that has remained buried for thousands of years, says Guggenberger.

Critics of the grazing restrictions in Tibet say that the government has applied them in a blanket way, without proper study and without



taking on board scientific findings. In some cases, they make sense, says Tsechoe Dorji, an ecologist at the ITPR’s Lhasa branch, who grew up in a herder family in western Tibet. “A total grazing ban can be justified in regions that are severely degraded”, he says, but he objects to the simple system used by the government to classify the health of the grasslands. It only considers the percentage of land covered by vegetation and uses the same threshold for all areas, without adjusting for elevation or natural moisture levels.

“Pastures with 20% vegetation cover, for instance, could be severely degraded at one place but totally normal at another,” says Dorji. This means that some of the grasslands that are classified as severely degraded are actually doing fine — and the grazing ban is actually hurting the ecosystem. “Having a sweeping grazing policy regardless of geographical variations is a recipe for disasters,” he says.

FAST FORWARD

China’s grazing policy is only one of several factors responsible for such damaging changes, say the researchers. Pollution, global warming and a rash of road-building and other infrastructure-construction projects have all taken a toll on the grasslands.

Ten days after leaving Xining, we caught a glimpse of Tibet’s future when we arrived at Nam Tso, a massive glacial lake in the southern part of the plateau. Here Dorji and Kelly Hopping, a graduate student at Colorado State University in Fort Collins, have been turning the clock forward by surrounding small patches of grassland with open-topped plastic chambers that artificially raise the temperature. These experiments are important because Tibet is a hotspot in terms of climate change; the average temperature on the plateau has soared by 0.3–0.4 °C per decade since 1960 — about twice the global average.

In trials over the past six years, they found that *Kobresia pygmaea*, the dominant sedge species, develops fewer flowers and blooms much later under warming conditions⁶. Such changes, says Dorji, “may compromise its reproductive success and long-term competitiveness”.

At the experimental site, the artificially warmed pastures have been taken over by shrubs, lichens, toxic weeds and non-palatable grass species, says Hopping. But when the researchers added snow



Near the headwaters of the Yellow River, lush grasslands have given way to sand dunes.



across the plateau and re-evaluate whether individual regions are degraded enough to merit a ban on grazing. “Unless the pastures are severely degraded, moderate grazing will help to restore the ecosystems,” he says.

But scientists are not banking on such reforms happening soon. Policies in Tibet are driven less by scientific evidence than by bureaucrats’ quest for power and funds, says a Lhasa-based researcher who requests anonymity for fear of political repercussions. Local officials often lobby Beijing for big investments and expensive projects in the name of *weiwén* (meaning ‘maintaining stability’). Because resistance to Chinese control over Tibet continues to flare up, the government is mostly concerned with maintaining political stability, and it does not require local officials to back up plans with scientific support, says the researcher. “As long as it’s for *weiwén*, anything goes.”

But officials such as Guo say that their policies are intended to help Tibet. “Although there is certainly room for improvement in some of the policies, our primary goals are to promote economic development and protect the environment,” he says.

Far away from Lhasa, herders such as Dodra say that they are not seeing the benefits of government policies. After we finish our visit at his home, Dodra’s entire family walks us into the courtyard — his mother-in-law spinning a prayer wheel and his children trailing behind. It has stopped snowing, and the sky has turned a crystal-clear, cobalt blue. “The land has served us well for generations,” says Dodra as he looks uneasily over his pasture. “Now things are falling apart — but we don’t get a say about how best to safeguard our land and future.” ■

Jane Qiu is a freelance writer in Beijing. Her trip across the Tibetan Plateau was supported by the SciDev.Net Investigative Science Journalism Fellowship for the Global South.

1. Cao, J., Yeh, E. T., Holden, N. M., Yang, Y. & Du, G. J. *Arid Environ.* **97**, 3–8 (2013).
2. Shen, M. et al. *Proc. Natl Acad. Sci. USA* **112**, 9299–9304 (2015).
3. Miehe, G. et al. *Quat. Sci. Rev.* **86**, 190–209 (2014).
4. Lu, H., Wang, S. S., Zhou, Q. W., Zhao, Y. N. & Zhao, B. Y. *Rangeland J.* **34**, 329–339 (2012).
5. Hafner, S. et al. *Glob. Chang. Biol.* **18**, 528–538 (2012).
6. Dorji, T. et al. *Glob. Chang. Biol.* **19**, 459–472 (2013).

to some heated plots, *Kobresia* did not lose out to the other plants, which suggests that the loss of soil moisture might be driving the shift in species. Higher temperatures increase evaporation, which can be especially potent at high elevations. “This is not good news for species with shallow roots”, such as the *Kobresia* favoured by livestock, she says.

Piao says that “this interplay between temperature and precipitation illustrates the complexity of ecosystem responses to climate change”. But researchers have too little information at this point to build models that can reliably predict how global warming will affect the grasslands, he says. To fill that gap, Wang and his colleagues started a decade-long experiment in 2013 at Nagchu, where they are using heat lamps to warm patches of grassland by precise amounts, ranging from 0.5°C to 4°C. They are also varying the amount of rainfall on the plots, and they are measuring a host of factors, such as plant growth, vegetation composition, nutrient cycling and soil carbon content. They hope to improve projections for how the grasslands will change — and also to determine whether there is a tipping point that would lead to an irreversible collapse of the ecosystem, says Piao.

PLATEAU PROGNOSIS

A fortnight into the trip, we finally arrived at the outskirts of Lhasa. At the end of the day, herders were rounding up their sheep and yaks in the shadows cast by snow-capped peaks. They and the other pastoralists across the plateau will have a difficult time in coming decades, says Nyima. Climate change was not a consideration when grassland policies were conceived over a decade ago, and so “many pastoralists are ill prepared for a changing environment”, he says. “There is a pressing need to take this into account and identify sound adaptation strategies.”

As a start, researchers would like to conduct a comprehensive survey of plant cover and vegetation composition at key locations across different climate regimes. “The information would form the baseline against which future changes can be measured,” says Wang. Many scientists would also support changes to the grazing ban and fencing policies that have harmed the grasslands. Dorji says that the government should drop the simplistic practice of ‘one policy fits all’



BOREDOM GETS INTERESTING

Implicated in everything from traumatic brain injury to learning ability, boredom turns out to be anything but boring.

BY MAGGIE KOERTH-BAKER

In 1990, when James Danckert was 18, his older brother Paul crashed his car into a tree. He was pulled from the wreckage with multiple injuries, including head trauma.

The recovery proved difficult. Paul had been a drummer, but even after a broken wrist had healed, drumming no longer made him happy. Over and over, Danckert remembers, Paul complained bitterly that he was just — bored. “There was no hint of apathy about it at all,” says Danckert. “It was deeply frustrating and unsatisfying for him to be deeply bored by things he used to love.”

A few years later, when Danckert was training to become a clinical neuropsychologist, he found himself working with about 20 young men who had also suffered traumatic brain injury. Thinking of his brother, he asked them whether they, too, got bored more easily than they had before. “And every single one of them,” he says, “said yes.”

Those experiences helped to launch Danckert on his current research path. Now a cognitive neuroscientist at the University of Waterloo in Canada, he is one of a small but growing number of investigators

ILLUSTRATION BY PATRYCJA PODKOŚCIELNY

engaged in a serious scientific study of boredom.

There is no universally accepted definition of boredom. But whatever it is, researchers argue, it is not simply another name for depression or apathy. It seems to be a specific mental state that people find unpleasant — a lack of stimulation that leaves them craving relief, with a host of behavioural, medical and social consequences.

In studies of binge-eating, for example, boredom is one of the most frequent triggers, along with feelings of depression and anxiety^{1,2}. In a study of distractibility using a driving simulator, people prone to boredom typically drove at higher speeds than other participants, took longer to respond to unexpected hazards and drifted more frequently over the centre line³. And in a 2003 survey, US teenagers who said that they were often bored were 50% more likely than their less-frequently bored peers to later take up smoking, drinking and illegal drugs⁴.

Boredom even accounts for about 25% of variation in student achievement, says Jennifer Vogel-Walcott, a developmental psychologist at the Cognitive Performance Group, a consulting firm in Orlando, Florida. That's about the same percentage as is attributed to innate intelligence. Boredom is "something that requires significant consideration", she says.

Researchers hope to turn such hints into a deep understanding of what boredom is, how it manifests in the brain and how it relates to factors such as self-control. But "it's a ways out before we're answering those questions", says Shane Bench, a psychologist who studies boredom in the lab of Heather Lench at Texas A&M University in College Station. In particular, investigators need better ways to measure boredom and more reliable techniques for making research subjects feel bored in the lab.

Still, the field is growing. In May 2015, the University of Warsaw drew almost 50 participants to its second annual conference on boredom, which attracted international speakers from social psychology and sociology. And in November, Danckert brought together about a dozen investigators from Canada and the United States for a workshop on the subject.

Researchers in fields from genetics to philosophy, psychology and history are starting to work together on boredom research, says John Eastwood, a psychologist at York University in Toronto, Canada. "A critical mass of people addressing similar issues creates more momentum."

A MEASURE OF MALAISE

The scientific study of boredom dates back to at least 1885, when the British polymath Francis Galton published⁵ a short note in *Nature* on 'The Measure of Fidget' — his account of how restless audience members behaved during a scientific meeting. But decades passed with only a few people taking a serious interest in the subject. "There are things all around us that we don't think to look at, maybe because they appear trivial," says Eastwood.

That began to change in 1986, when Norman Sundberg and Richard Farmer of the University of Oregon in Eugene published their Boredom Proneness Scale (BPS)⁶, the first systematic way for researchers to measure boredom — beyond asking study participants, "Do you feel bored?". Instead, they could ask how much participants agreed or disagreed with statements such as: "Time always seems to be passing slowly"; "I feel that I am working below my abilities most of the time" and "I find it easy to entertain myself". (The statements came from interviews and surveys that Sundberg and Farmer had conducted on how people felt when they were bored.) A participant's aggregate score would give a measure of his or her propensity for boredom.

The BPS opened up new avenues of research and made it apparent that boredom was about restlessness as much as apathy, the search for meaning as much as ennui. It has served as a launching point for other boredom scales, a catalyst for making the field more important and a tool for connecting boredom to other factors, including mental health and academic success.

But it also has some widely acknowledged flaws, says Eastwood. One is that the BPS is a self-reported measure, which means that it is inherently subjective. Another is that it measures susceptibility to boredom — 'trait boredom' — not

the intensity of the feeling in any given situation, which is known as state boredom. Studies consistently show that these two measures are independent of each other, yet researchers are only beginning to tease them apart.

This can be particularly confounding in educational settings. Shifts in teaching style or classroom environment are unlikely to reduce students' trait boredom, which is intrinsic and slow to change, but can be very effective at reducing state boredom, which is purely situational. The BPS has often been misused to measure both forms of boredom at the same time, yielding answers that are likely to be misleading, says Eastwood.

Scientists are still hashing out how to improve on the BPS. In 2013,

WHEN SHE WAS WORKING ON HER DISSERTATION, SHE DECIDED TO MAKE A VIDEO THAT WOULD BORE MOST PEOPLE TO TEARS.

Eastwood helped to develop the Multidimensional State Boredom Scale (MSBS)⁷, which features 29 statements about immediate feelings, such as: "I am stuck in a situation that I feel is irrelevant." Unlike the BPS, which is all about the participant's habits and personality, the MSBS attempts to measure how bored people feel in the moment. And that, Eastman hopes, will give it a better shot at revealing what boredom is for everybody.

But to measure boredom, researchers must first make sure that study participants are bored. And that is a whole different challenge.

THE MOST BORING VIDEO EVER

One way to create a particular mood, used for decades in psychology, is to show people a video clip. There are scientifically validated videos for inducing happiness, sadness, anger, empathy and many other emotions. So when she was working on her dissertation at Waterloo in 2014, Colleen Merrifield decided to make a video that would bore most people to tears.

In Merrifield's video, two men stand in a white, windowless room. Silently, they take clothes from a pile between them and hang them on a white rack — a camisole, a shirt, a sweater, a sock. The seconds tick by: 15, 20, 45, 60. The men keep hanging laundry. Eighty seconds. One of the men asks the other for a clothes peg. One hundred seconds. They keep hanging laundry. Two hundred seconds. They keep hanging laundry. Three hundred seconds. They keep hanging laundry. Shown on a loop, the video can last for as long as five and a half minutes.

Perhaps unsurprisingly, the people to whom Merrifield showed this found it stupefyingly dull⁸. But then she tried using the video to study how boredom affected the ability to focus and pay attention. Her protocol called for participants to carry out a classic cognitive attention task — watching for star-like light clusters to appear or disappear on a monitor — then to sit through the video to get good and bored, and finally to do the task again so that she could see how boredom affected their performance. But she found that she had to redesign the experiment: the task was boring people more than the video.

This was not entirely unexpected. Previous studies of boredom had often used tasks instead of videos. But it also demonstrated the problem. There are so many ways for researchers to bore people with tasks — asking them to proofread address labels, say, or to screw nuts and bolts together — that it had always been difficult to compare individual studies. For instance, different studies have found boredom to be correlated with both rising and falling heart rate⁹. But without a standardized method for inducing boredom, it is impossible to work out who is right.

In 2014, researchers at Carnegie Mellon University in Pittsburgh, Pennsylvania, published a paper⁹ that aimed to begin the process of standardization. It compared six different boredom inductions, representing three broad classes — repetitive physical tasks, simple

➔ **NATURE.COM**
To find your score
on the Boredom
Proneness Scale, see:
go.nature.com/xpgwok

cognitive tasks, and video or audio media — as well as a control video. The researchers used the MSBS to see how intensely each task elicited boredom, and a measure called the Differential Emotion Scale to see whether each task elicited boredom alone, or a number of other emotions. All six tasks were significantly more boring than the control and all six caused boredom almost exclusively. The best of the bunch was a task that required participants to click a mouse button to rotate a computer icon of a peg a quarter of a turn clockwise, over and over.

After that, says Danckert, “I think I might be abandoning the video” to induce boredom in the lab. Instead, he will rely on behavioural tasks.

The inexactness of the tools leaves holes in what researchers can reasonably say about boredom. For instance, many real-world problems that are highly correlated with boredom are connected to the idea of self-control, including addiction, gambling and binge-eating¹⁰. “I characterize boredom as a deficiency in self-regulation,” Danckert says. “It’s a difficulty of engaging with tasks in your environment. The more self-control you have, the less likely you are to be bored.”

But does this mean that self-control and boredom are measures of the same thing? Even Danckert is uncertain. Consider people with a history of traumatic brain injury. “Failures of self-control are their problem,” he says. “They might be inappropriately impulsive; there’s increased risk-taking; they might also engage in drug and alcohol abuse,” Danckert certainly saw his brother, Paul, experience all those things in the wake of his injury.

But in Danckert’s research sample of people with traumatic brain injury — who are predominantly in their 40s — ageing seems to have weakened the link between boredom and self-control. In data that are not yet published, Danckert says, his patients report levels of self-control no lower than those of the general population, but their boredom-proneness scores are much higher. By contrast, Danckert’s brother seems to demonstrate the opposite effect. He struggled for years with self-control issues, but eventually became less bored and reclaimed his love of music. “It’s the most important thing in his life, next to his children,” Danckert says.

So there is reason to suspect that boredom and self-control can exist independently — but there is not yet enough evidence to understand much beyond that.

PAINFULLY DULL

Despite all this uncertainty, researchers see themselves as laying a foundation, creating tools and standards that will allow them to tackle really important questions. “We’re establishing boredom as a testable construct,” says Bench.

Defining boredom is an important part of that. Different researchers have different pet definitions: a German-led team, for example, identifies five types of boredom¹¹. But most workers in the field agree that, at least some of the time, people will work very hard to relieve boredom. This not only presents a more active version of boredom than most people are probably used to, but also has tangible connections to efforts to address boredom in the real world.

Lench and Bench are testing whether the drive to become un-bored is so strong that people might be willing to choose unpleasant experiences as an alternative. This idea builds on research that has shown a correlation between sensation-seeking behaviour, even risky behaviour, and high boredom-proneness scores¹². It is also similar to findings published in *Science*¹³ in 2014 and *Appetite*¹⁴ in 2015. In the first study, researchers asked people to sit in a room with nothing to do for as long as 15 minutes at a time. Some of the participants, particularly men, were willing to give themselves small electric shocks rather than be left alone with their thoughts. The second paper described two experiments: one in which the participants had access to unlimited sweets, and another in which they had access to unlimited electric shocks. Participants ate more when they were bored — but they also gave themselves more shocks. Even when it is not very pleasant, apparently, novelty is better than monotony.

Novelty might also have a role in overcoming boredom in the classroom. In 2014, for instance, researchers led by psychologist Reinhard Pekrun of the University of Munich in Germany reported¹⁵ how they had followed 424 university students over the course of an

academic year, measuring their boredom levels and documenting their test scores. The team found evidence of a cycle in which boredom begot lower exam results, which resulted in more disengagement from class and higher levels of boredom. Those effects were consistent throughout the school year, even after accounting for students’ gender, age, interest in the subject, intrinsic motivation and previous achievement. But other studies suggest that novelty can disrupt this cycle¹⁶.

Sae Schatz, director of the Advanced Distributed Learning Initiative, a virtual company that develops educational tools for the US Department of Defense, points to one experiment¹⁷ with a computer system that tutored students in physics. When the system was programmed to insult those who got questions wrong and snidely praise those who got them right, says Schatz, some students, especially adult learners, saw improved outcomes and were willing to spend longer on the machines. Schatz thinks that this could be because the insults provided enough

“WE’RE ESTABLISHING BOREDOM AS A TESTABLE CONSTRUCT.”

novelty to keep people engaged and less prone to boredom.

Looking to the future, researchers such as Eastwood are intent on finding better ways to understand what boredom is and why it is correlated to so many other mental states. They also want to investigate boredom in people who aren’t North American college students. That means testing older people, as well as individuals from diverse ethnic and national backgrounds. And, given the impact that boredom may have on education, it also means developing versions of the BPS and MSBS that can be administered to children.

Many researchers likewise hope to expand on the types of study being done. To get beyond self-reported data, Danckert wants to start looking at brain structures, and seeing whether there are differences between people who score highly on the BPS and those who don’t. These data could help him to understand why boredom manifests so strongly in some people with traumatic brain injury.

There’s also a need, Danckert says, for more scientists to realize that boredom is fascinating. “We may be on the cusp of having enough people to advance a little more quickly,” he says. ■

Maggie Koerth-Baker is a freelance writer in Minneapolis, Minnesota.

1. Ganley, R. M. *Int. J. Eat. Disorders* **8**, 343–361 (1989).
2. Vanderlinden, J., Grave, R. D., Vandereycken, W. & Noorduin, C. *Eat. Behav.* **2**, 79–83 (2001).
3. Kass, S. J., Beede, K. E. & Vodanovich, S. J. *Accident Anal. Prev.* **42**, 874–880 (2010).
4. The National Center on Addiction and Substance Abuse at Columbia University. *National Survey of American Attitudes on Substance Abuse VIII: Teens and Parents* (Columbia Univ., 2003); available at <http://go.nature.com/vcohao>
5. Galton, F. *Nature* **32**, 174–175 (1885).
6. Farmer, R. & Sundberg, N. D. *J. Personal. Assess.* **50**, 4–17 (1986).
7. Fahlman, S. A., Mercer-Lynn, K. B., Flora, D. B. & Eastwood, J. D. *Assessment* **20**, 68–85 (2013).
8. Merrifield, C. *Toward a Model of Boredom: Investigating the Psychophysiological, Cognitive, and Neural Correlates of Boredom* PhD thesis, Univ. Waterloo (2014).
9. Markey, A., Chin, A., Vanepps, E. M. & Loewenstein, G. *Percept. Motor Skills* **119**, 237–253 (2014).
10. Struk, A. A., Scholer, A. A. & Danckert, J. *Cognition Emotion* <http://dx.doi.org/10.1080/02699931.2015.1064363> (2015).
11. Goetz, T. et al. *Motiv. Emotion* **38**, 401–419 (2014).
12. Elpidourou, A. *Front. Psychol.* **5**, 1245 (2014).
13. Wilson, T. D. et al. *Science* **345**, 75–77 (2014).
14. Havermans, R. C., Vancleef, L., Kalamatanios, A. & Nederkoorn, C. *Appetite* **85**, 52–57 (2015).
15. Pekrun, R., Hall, N. C., Goetz, T. & Perry, R. P. *J. Educ. Psychol.* **106**, 696–710 (2014).
16. Vogel-Walcutt, J. J., Fiorella, L., Carper, T. & Schatz, S. *Educ. Psychol. Rev.* **24**, 89–111 (2012).
17. Graesser, A. C. *Am. Psychol.* **66**, 746–757 (2011).

COMMENT

ENTOMOLOGY Biography of a notorious moth and mosquito man **p.152**



AGEING With lifespan increasing, how will we extend healthspan? **p.154**

ENERGY Inspect ageing fuel pipelines, they could fail at any time **p.156**

CONSERVATION Europe's refugee fences will isolate and entangle wildlife **p.156**

JIM WEST/ALAMY



Drums containing contaminated materials from the US nuclear-defence programme are stored at the Waste Isolation Pilot Plant in New Mexico.

Reassess New Mexico's nuclear-waste repository

Proposals to bury plutonium from nuclear weapons must address chemical interactions and intrusion risks, say **Cameron L. Tracy, Megan K. Dustin and Rodney C. Ewing**.

More than 600 metres below ground near Carlsbad, New Mexico, is the world's only operating deep geological repository currently accepting transuranic nuclear waste: that contaminated by elements heavier than uranium. The Waste Isolation Pilot Plant (WIPP), run by the US Department of Energy (DOE), is used to dispose of laboratory equipment, clothing and residues from the nation's nuclear-defence programme. In the past 15 years, around 91,000 cubic metres (equivalent to covering a soccer field to a depth of about 13 metres)

of such transuranic waste, mostly of relatively low radiation levels, has been placed there.

The main contaminants are long-lived isotopes of plutonium (mainly plutonium-239, with a half-life of 24,100 years, and plutonium-240, with a half-life of 6,560 years) and shorter-lived isotopes of americium and curium. In rooms carved out of a 250-million-year-old salt bed, the waste is stored in hundreds of thousands of plastic-lined steel drums. The repository is now at about half of its planned capacity and is to be sealed in 2033.

The DOE is responsible for performing safety assessments to ensure that WIPP will not exceed limits on exposure to radioactivity, as set by the US Environmental Protection Agency (EPA), for 10,000 years.

But new demands are emerging. An arms-control agreement with Russia made in 2000 obliges the United States to dispose of 34 tonnes of plutonium from dismantled nuclear weapons¹. Following the terms of the agreement, the United States planned to convert the material into a fuel — mixed (uranium and plutonium) oxide, or MOX ►

► — to burn in commercial nuclear-power plants. But faced with soaring construction costs for a MOX fabrication facility at the Savannah River Site in South Carolina, the DOE has commissioned evaluations of alternatives².

The most recent report³, published in August 2015, recommends burying the weapons' plutonium at WIPP. Judging the repository's performance to have been "successfully demonstrated", the DOE's Red Team expert panel proposes that the 34 tonnes of weapons plutonium can be added to WIPP once it has been diluted to low concentrations comparable to that of the transuranic waste at WIPP.

In fact, WIPP's safety record is mixed. On 14 February 2014, a burst drum released small quantities of plutonium and americium to the surface (with a radioactivity of around 100 millicuries, or 3.7 gigabecquerels)⁴. Airborne radioactive material reached the surface through the ventilation system and spread 900 metres from the repository's exhaust shaft. Twenty-one workers were exposed to low levels of radioactivity, the highest dose equivalent to that from a chest X-ray. Nine days earlier, smoke from a burning truck filled the underground workings and shaft, damaging mechanical, electrical and ventilation systems.

The DOE says that such accidents do not compromise the long-term performance of the repository. We agree that they need not — if lessons are learned. Our concern is not the events' severity but that they were unanticipated. These accidents illustrate how difficult it is to predict potential failures of such a disposal system over millennia. For example, assumptions about the repository's geochemistry or the likelihood of drilling into it can lead to underestimates of the risks.

Before expanding WIPP's plutonium inventory, the DOE must examine more carefully its safety assessment for performance that stretches to 10,000 years and beyond.

CULTURE OF COMPLACENCY

The 2014 radioactive leak at WIPP was caused by heat from a chemical reaction in a drum⁴. Plutonium-contaminated nitrate salts, a waste product of plutonium purification at Los Alamos National Laboratory (LANL) in New Mexico, reacted with an organic, wheat-based commercial cat litter used as an absorbent for liquid wastes. The heat popped the lid. Although sensors detected the released radioactivity and diverted exhaust air through filters, some radioactive material leaked through. WIPP operators sealed the leak in the filtration system and sealed off the room in which the leak occurred. The breached drum remains in the repository.

Analyses of the accidents⁴ by the DOE have documented a lack of a 'safety culture' at WIPP. The facility's successful operation for 15 years had bred complacency. The failures

were wide-ranging: in safety assessments, control of drum contents, installation and maintenance of equipment, and preparation for an accident. An investigation of the drum-packaging procedure, for example, found "no evidence that any type of technical evaluation occurred" when selecting the organic absorbent material, even though its incompatibility with nitrate salts had been raised at LANL during waste packaging⁴.

From a systems-analysis perspective, the drum breach was a 'normal' accident⁵ — a human mistake that led to a cascade of errors and breakdowns, exacerbated by a failure to enforce safety protocols. Complex technologies are prone to unanticipated failures that can progress quickly; examples include the 1979 Three Mile Island nuclear-plant meltdown in Pennsylvania and the 1986 Challenger space-shuttle explosion. Such accidents cannot be easily predicted, but a system designed with failure in mind can mitigate the risk.

The WIPP accident can be taken as a positive — it presents an opportunity to learn. The DOE has aggressively identified its causes and implemented corrective actions; incompatible chemicals are no longer mixed in the drums. But once the repository is closed, its contents cannot be monitored or problems fixed. We cannot be certain that future inhabitants of the area will even know that WIPP is there. To put the timescales in perspective, agriculture was developed just over 10,000 years ago.

LONG-TERM SAFETY

WIPP's present safety assessment addresses two scenarios: first, undisturbed performance and, second, human intrusion, such as inadvertently drilling through the repository in search of oil and gas⁶. The first foresees that after closure, the salt into which the repository is built will deform and flow around the drums to encase the waste. The model assumes that no fluids, such as brine, are present and that the site remains geologically isolated. Although the drums will be crushed, the radioactive material will be locked in the dry, solid salt, with no way to release radioactivity to the biosphere. Reliance on the geological barrier is so great that the form and composition of the waste is assumed to be unimportant; it need not even be treated.

Human intrusion could release radioactivity to the environment⁶. Salt deposits, layered as sediments or as salt domes, are often associated with mineral and energy resources, such as potash and hydrocarbons — oil and gas. In southeastern New Mexico, exploration for and extraction of these fuels has led to extensive drilling in the Permian

Basin, where WIPP is located.

The probability of a borehole piercing the repository in the next 10,000 years is significant. If a borehole were to puncture the repository and a brine pocket, which are known to exist in the Castile geological formation below the Salado salt formation in which the repository sits, fluid may reach the transuranic waste (see 'Accident risk'). To assess the risk of radioactive release, one must first establish the probability of borehole penetration and determine how the pressurized brine will react with the waste.

In forecasting future drilling rates, the EPA has used a 100-year historical average rate for the region, which predicts 67.3 boreholes per square kilometre over the 10,000-year regulated period⁶. But drilling near WIPP has risen sharply in recent years. As horizontal drilling and hydraulic-fracturing techniques have made new areas of hydrocarbon-bearing rocks accessible, the Permian Basin has become the most prolific oil-producing area in the United States. A recent 10-year historical average (2002–12) yields 148 boreholes per square kilometre over 10,000 years, more than doubling the projected risk of repository intrusion. Drilling rates, the effects of new technologies, and supply and demand pressures on hydrocarbon production are difficult if not impossible to predict centuries ahead.

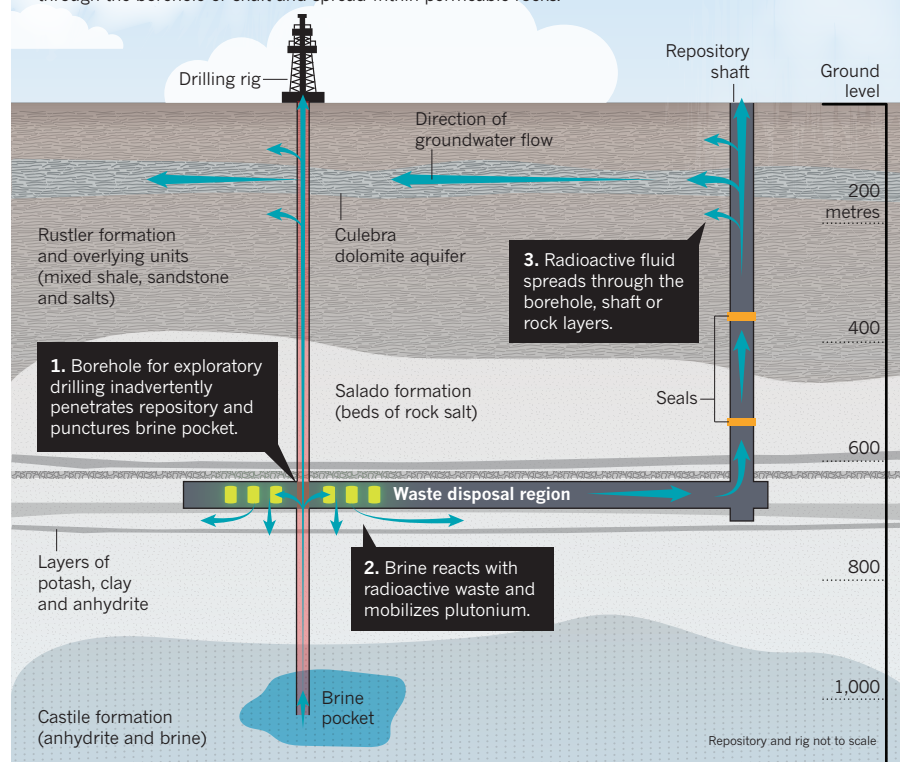
The concentration of transuranic elements leached by intruding brine is also hard to estimate because of the complexity of the waste: a typical drum contains a variety of materials, such as lab coats, gloves and other laboratory equipment. Different micro-geochemical environments will develop around different waste types. Chemically organic materials, such as plastic bags, may degrade by microbial action and generate carbon dioxide. In brine, CO₂ forms stable carbonate and bicarbonate complexes with plutonium and other actinides (elements 89–103), raising their concentrations in solution. Large bags of magnesium oxide powder, amounting to more than 31,000 tonnes, have been placed in WIPP disposal rooms as an 'engineered barrier'. The magnesium oxide should react with the CO₂ to form stable magnesium carbonates, thereby removing CO₂ from solution and reducing the solubility of actinides. This presumes that the reactions proceed to completion and all the CO₂ is consumed.

The safety analysis calculations for WIPP assume that there is no CO₂ present, dramatically lowering actinide concentrations in the brine and thus the risk of release of radioactivity. But reliance on magnesium oxide and a series of idealized reactions to constrain the repository's geochemistry is problematic, particularly if the amount of plutonium stored at WIPP increases. As made clear by the 2014 accidents, complex interactions of materials must be carefully considered when predicting the repository's

"We cannot be certain that future inhabitants of the area will even know WIPP is there."

ACCIDENT RISK

Thousands of years in the future, inadvertently drilling a borehole through the Waste Isolation Pilot Plant, a nuclear-waste repository, and into a brine pocket could release radioactive material into the environment. The brine would interact with the waste and contaminated fluid could reach the surface through the borehole or shaft and spread within permeable rocks.



performance now and in the future.

The Red Team report proposes diluting the weapons plutonium before its disposal in an “inert adulterant” — a classified mixture of cementing, gelling, thickening and foaming agents known as stardust. The report is unclear on what is meant by ‘inert’; however, inert materials are rare, particularly those that must remain so for thousands of years.

PLUTONIUM DISPOSAL

In the case of plutonium-bearing solids, demonstrating chemical inertness presents a huge challenge. In near-surface conditions, plutonium can assume a variety of oxidation states — up to four, each with different solid-state and geochemical behaviours⁷. Its decay product uranium-235 has two principal oxidation states, U^{4+} and U^{6+} , each with different geochemical mobility⁷. This complexity makes it difficult to predict how the actinides will react or be transported.

Also, actinides decay mainly by the emission of α particles (energetic helium nuclei). During each decay, the daughter nucleus recoils and displaces thousands of atoms in the surrounding solid. Over time, this damage accumulates and changes the properties and chemical stability of the material. Radiation effects in actinide-bearing materials have been well documented over the past 20 years⁸, but are not considered in the Red

Team’s evaluation.

The ‘dilute-and-dispose’ proposal to convert weapons-plutonium pits to plutonium oxide for burial in WIPP³ immediately raises safety issues. The extra plutonium nearly triples the current projected plutonium (around 12 tonnes) at closure. The design and safety assessment did not envision such a large amount. WIPP’s capacity would have to expand by 15%, increasing the likelihood that a borehole will one day intersect it.

And the changed inventory of actinides demands new assessments of interactions with the materials present, including brine and CO_2 . The amount of plutonium mobilized in brine depends on its solubility, which depends on its form and the amount of CO_2 present after reaction with the bags of magnesium oxide.

NEXT STEPS

The current regulatory period of 10,000 years is short relative to the 24,100-year half-life of plutonium-239, let alone that of uranium-235, which has a half-life of 700 million years. To accommodate the extra plutonium, the regulatory period might be lengthened, meaning that the probability of human intrusion during this period increases.

Some of these issues and others were raised in two 2015 reviews^{9,10} of the Red Team report by the consultancy High Bridge Associates of

Greensboro, Georgia. But the analysis did not consider the possibility of human intrusion.

WIPP is fulfilling an important national need — the disposal of legacy transuranic waste from US defence programmes. Its opening was the culmination of 20 years of scientific research, engineering design and public engagement. Despite the accidents, WIPP can still fulfil its mission.

However, proposals to substantially increase the plutonium inventory combined with a failure to revise the safety assessment, particularly the possibility of human intrusion, bear witness to the ease with which policy decisions can disregard the fundamental science — and risk yet another failure.

The Red Team report shows a limited effort to consider or manage inherent risks. The shortcomings of proposals to dispose of weapons plutonium at WIPP mirror the operational failings that led to the 2014 accidents. Before the DOE considers implementing these recommendations, it should look to the repository’s record over the past 15 years of operation and reassess its confidence in the safe performance of the facility over the next 10,000. ■

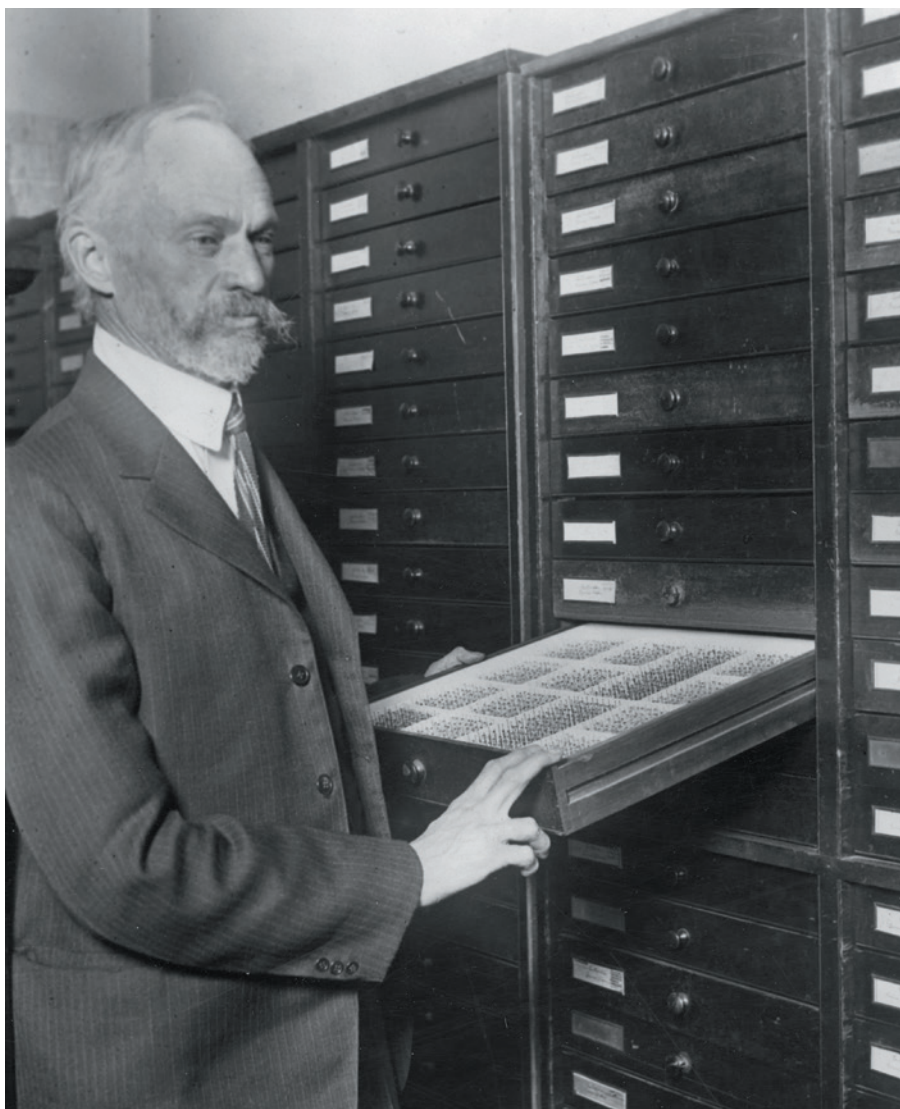
Cameron L. Tracy is a MacArthur postdoctoral fellow at the Center for International Security and Cooperation, Stanford University, California, USA.

Megan K. Dustin is a graduate fellow in the Department of Geological Sciences, Stanford University, California, USA. **Rodney C.**

Ewing is Frank Stanton professor in nuclear security at the Center for International Security and Cooperation, Stanford University, California, USA.
e-mail: rewing1@stanford.edu

1. Clements, T., Lyman, E. & von Hippel, F. *Arms Control Today* **43**, 8–15 (2013).
2. US Department of Energy. *Report of the Plutonium Disposition Working Group: Analysis of Surplus Weapons-Grade Plutonium Disposition Options* (DOE, 2014).
3. US Department of Energy. *Final Report of the Plutonium Disposition Red Team* (DOE, 2015).
4. US Department of Energy. *Accident Investigation Report: Phase 2, Radiological Release Event at the Waste Isolation Pilot Plant, February 14, 2014* (DOE, 2015).
5. Perrow, C. *Normal Accidents: Living with High Risk Technologies* (Princeton Univ. Press, 1999).
6. US Department of Energy. *Compliance Recertification Application 2014 for the Waste Isolation Pilot Plant* (DOE, 2014).
7. Maher, K., Bargar, J. R. & Brown, G. E. *Inorg. Chem.* **52**, 3510–3532 (2012).
8. Ewing, R. C. & Weber, W. J. in *The Chemistry of the Actinide and Transactinide Elements* Vol. 6 3813–3888 (eds Morss, L. R., Edelstein, N. M., & Fuger, J.) (Springer, 2010).
9. High Bridge Associates. *Plutonium Disposition Study Options* (High Bridge Associates, 2015).
10. High Bridge Associates. *Supplemental Report, Independent Assessment of the Impact of Disposing of Surplus Weapons Plutonium at WIPP* (High Bridge Associates, 2015).

The views expressed are those of the authors and not those of the US Nuclear Waste Technical Review Board, of which R.C.E. is a member.



Harrison Dyar collected 500,000 different kinds of mosquitoes over his career.

ENTOMOLOGY

A life of insects and ire

Professional feuds and private oddities abound in a biography of Harrison Dyar, finds **William Foster**.

Harrison G. Dyar Jr (1866–1929) was an influential US biologist who became notorious in his own lifetime. His pioneering work changed our understanding of the biology and systematics of two globally important groups: the Lepidoptera (butterflies and moths) and mosquitoes. But his scientific legacy is overshadowed by his protracted, spectacularly

belligent feuds with fellow entomologists, and scandalous revelations about his private life. For some 14 years, he was married to two women, maintaining two families of five children in all. Towards the end of his life, he built an extensive system of brick-lined tunnels deep under the heart of Washington DC. When accidentally discovered in 1924, the tunnels caused such a sensation

Moths, Myths, and Mosquitoes: The Eccentric Life of Harrison Dyar

MARC EPSTEIN
Oxford University Press: 2016.

that they featured in cartoons in that year's presidential campaign.

Insect taxonomist Marc Epstein has a long association with

the entomology department of the Smithsonian Institution in Washington DC, where Dyar spent most of his professional career. To write *Moths, Myths, and Mosquitoes*, Epstein dug deep into the Smithsonian's holdings of publications, correspondence, diaries, unpublished short stories and novellas, newspaper articles and marriage certificates, providing rich context for Dyar's intellectual and scientific milieu. Epstein's descriptions of Dyar's collecting trips and battles with colleagues are particularly evocative.

Dyar's major scientific achievement was to establish the use of insects' immature stages in the construction of evolutionary trees, or phylogenies, based on Darwinian principles. His 1902 list of North American Lepidoptera has stood the test of time, as have most of his species classifications among the Lepidoptera, sawflies and mosquitoes. Among biologists he is perhaps best known for formulating Dyar's Law, which postulates that the width of the head capsule of each developmental stage, or instar, of a caterpillar is related by a constant ratio to the width of the head capsule of the next instar. When looking at a sample, this takes most of the guesswork out of deciding which instar you are dealing with, whether any are missing and how many a particular species might have. Epstein shows that Dyar fudged his data a little to fit his model, but the model itself remains useful. (It certainly worked for me when I did my first biology project, a demonstration of Dyar's Law, as a schoolboy in the mid-1960s.)

It is relatively difficult to establish objective measures of accuracy in systematics and phylogeny; as a result, researchers are often tempted to settle arguments by sheer force of personality. Dyar embraced this temptation with rash enthusiasm, even though his own research had enhanced the objectivity of the field. He seemed to relish long-running scientific battles. In 1905, he accused Henry Skinner, editor of *Entomological News*, of indulging in a "hysterical outburst", adding, "Better take a sedative." Skinner responded that Dyar drank "too much ice water", so his blood was "too frigid" to understand what went on in the *News*. In 1908, mosquito expert Evelyn Mitchell sued Dyar for libel. The case was dismissed, and Dyar got his revenge by writing an unpublished short story, 'The Taming of a Suffragette', which features a thinly disguised Mitchell. As *Moths, Myths, and Mosquitoes*

NATURE.COM
Visit the Books & Arts blog:
nature.com/booksandarts

progresses, one begins to sense that the battles are what kept Dyar going. “Lest affairs become too monotonous, I feel obliged to start something occasionally,” he wrote in 1925, in a letter to Leland Howard, his superior at what is now the US National Museum of Natural History.

Much of the book, in particular the sections on Dyar’s private life, feels like the raw materials of a biography rather than a finished product. Epstein fields a wonderful set of characters, but struggles to breathe life into them; many passages are cluttered with detail. Nor is it always easy to work out the sequence of events, but the many excellent illustrations and photographs help to shape and colour the narrative. There is a particularly moving photo of Dyar and his two-year-old son Otis on a beach, staring rather gloomily into the camera. Next to them is a huge sand sculpture of a spiny limacodid caterpillar: even at the seaside, there is no escape from Dad’s work.

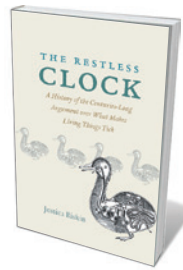
In the epilogue, Epstein offers definitive proof of when Dyar’s bigamy began. Dyar married Zella Peabody in 1889, fathering two children with her before their divorce in 1920. Yet in 1906, Dyar had secretly married Wellesca Pollock under the name of Wilfred P. Allen. He had three children with Pollock;

“He constructed a labyrinth of tunnels lined with bricks and furnished with electric lights.”

However appalling Dyar’s behaviour to relatives and colleagues (he described Smithsonian librarian Mathilde Carpenter as a “boisterous, screaming plebian” in a letter to her supervisor), there is something oddly admirable about his final obsession. While digging a hollyhock bed for Peabody, he found himself 2 metres down, and was seized with the urge to keep going. He constructed a labyrinth of tunnels lined with bricks and furnished with electric lights, sculpted animal heads and mottos from Virgil. He claimed that they provided him with exercise. And no biologist could fail to admire Dyar’s unquenchable respect for the insects that he studied. When the US government set out to exterminate mosquitoes in Yosemite National Park, California, Dyar protested, saying that a few bites are “good for hikers and lend zest to the fisherman’s waiting”. ■

William Foster is curator of insects at the University Museum of Zoology in Cambridge, UK; senior lecturer in the Department of Zoology at the University of Cambridge; and editor-in-chief of Biological Reviews.
e-mail: waf1@cam.ac.uk

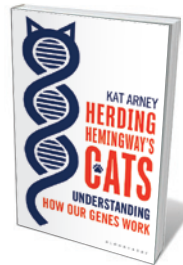
Books in brief



The Restless Clock: A History of the Centuries-Long Argument Over What Makes Living Things Tick

Jessica Riskin UNIVERSITY OF CHICAGO PRESS (2016)

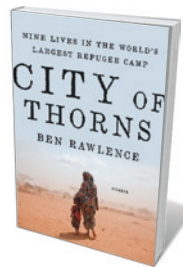
At the heart of this scientific and cultural history is the concept of agency — the capacity to act — in nature. Jessica Riskin reveals how two distinct interpretations emerged from the mechanical Universe of the Enlightenment: Isaac Newton’s passive version, reliant on a divine tinkerer; and Gottfried Leibniz’s, which saw life as purposeful and “self-transforming”. Riskin’s investigation of this duality, by way of Renaissance automatons, the gestation of evolutionary theory and quantum mechanics, is engrossing and illuminating.



Herding Hemingway’s Cats: Understanding How Our Genes Work

Kat Arney BLOOMSBURY SIGMA (2016)

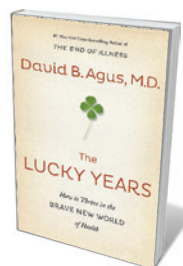
In this witty, clued-up report from the front lines of genetics, science communicator and broadcaster Kat Arney unravels the intricacies of the discipline with a romp through ‘thumbed’ cats, hipped fish and frank interviews with scientists such as evolutionary biologist Dan Graur. As she synthesizes key findings, she deploys a host of droll, yet apt, metaphors (including the human genome as a grim cable-television channel featuring tedious repeats), and pulls no punches in laying out the vast gaps in our understanding and the rancorous debates within the field.



City of Thorns: Nine Lives in the World’s Largest Refugee Camp

Ben Rawlence PICADOR (2016)

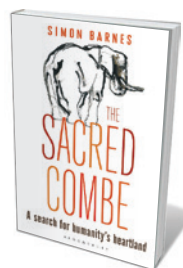
Dadaab in the Kenyan desert is the world’s largest refugee camp, a last-ditch home to some half a million people fleeing violence in the Horn of Africa. In this trenchant, densely layered sociopolitical study, investigative journalist Ben Rawlence reveals Dadaab’s complexities through the lives of nine residents, impossibly courageous survivors of derailed cultures, imploded cities and sundered families. A reminder that although there are thousands of refugees at Europe’s borders, millions more languish in camps — as Rawlence puts it, between “impossible dreams and a nightmarish reality”.



The Lucky Years: How to Thrive in the Brave New World of Health

David B. Agus SIMON & SCHUSTER (2016)

Oncologist and biomedical researcher David Agus’s bestselling *The End of Illness* (Simon & Schuster, 2012; see *Nature* **480**, 177; 2011) argued persuasively for personalized health care. In this clear-cut follow-up, he details health interventions including monitoring technologies, analysable aggregate data sets and, more controversially, smartphone apps that detect signs of depression. What is strongest here is Agus’s deft marshalling of research old and new, and his common-sense guidance on preventives such as sleep hygiene and the optimal level of exercise (450 minutes per week).



The Sacred Combe: A Search for Humanity’s Heartland

Simon Barnes BLOOMSBURY NATURAL HISTORY (2016)

The moment a herd of elephants ripped into his thatched hut did it for natural-history writer Simon Barnes: he suddenly realized that Zambia’s Luangwa Valley had claimed him for its own. This episodic journey into the wilds of Devonshire, Africa and memory — the edenic spaces where species fleetingly coexist — is studded with descriptive jewels. Here, for instance, are eland antelopes, “one-tonners drifting back like pale wisps of smoke”, and an otter with “elegant bum briefly sky-pointing” as it dives in for the hunt. [Barbara Kiser](#)

Extending the healthspan

Linda Partridge examines studies on preventative medicine for the ageing.

The chances of living to old age are higher than ever in many parts of the world. So, particularly in developed countries, health-care systems are struggling to cope with the 'silver tsunami' of elderly people with clusters of diseases for which age is the main risk factor, including cancer, diabetes, cardiovascular disease, sarcopenia and dementia. Fortunately, the opportunity is at hand to transform the landscape and keep people in better health as they age.

Aging is a collection of articles edited by gerontologists Jay Olshansky, George Martin and James Kirkland. It explores the potential to extend human health by drawing on discoveries about the biology of ageing. The overall coverage is US-centric, and would have benefited from more cutting-edge basic science and human demography from the rest of the world. But the message is clear: it is time to begin the revolution in medical approaches to ageing-related disease and late-life health.

In lab animals, *Aging* tells us, simple genetic and environmental interventions can increase healthy lifespan substantially. A restricted diet can protect ageing rodents and rhesus monkeys from most impairments and diseases. Genetic alterations to the signalling networks that sense and respond to nutrients and to other inputs can have similar effects. These interventions tamp down the changes that characterize ageing and that lead to pathology, including chronic inflammation, cellular senescence, damage to macromolecules and decline in stem-cell function.

An emerging consensus thus regards ageing as composed of modifiable sub-syndromes. These could be targeted with drug combinations and environmental changes to produce a broad-spectrum, preventive medicine for multiple ageing-related diseases. Currently, most treatment is directed towards individual diseases as they arise.

Development of new drugs to target ageing is a major challenge, thoughtfully addressed in the chapter by James Kirkland. Ageing is not recognized by regulatory authorities as a disease, or as a valid target of clinical trials, which would be prohibitively expensive because they would need to be of long duration, and to use initially disease-free participants. It would be more feasible to test drugs that have been approved for specific diseases, and that also target mechanisms of ageing, against other age-related



Woody, age 83.

conditions. For instance, older people tend to have a poor immune response to immunization against influenza; that response can be improved by pre-treating them with sirolimus, a drug already shown to increase animal lifespan and licensed as an immunosuppressant for use after kidney transplants. Polymorbidity (the presence of multiple conditions at once) has not yet been assessed as a potential drug target, and would require new, combinatorial measures of outcome and trials with older people, who are currently generally excluded from clinical trials. A 'poly pill' of drugs that target

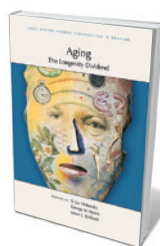
different axes of ageing may give maximum protection, posing further interesting challenges for trials.

Olshansky discusses the ethics of targeting ageing for disease prevention. Ageing, he shows, is not nature's way of making space for the young; rather, it is a haphazard process of decline. The discovery that it is malleable allows an entirely new approach to improving the health and welfare of older people. Lifespan might increase slightly as a result, but the crucial point is that morbidity could be compressed. Olshansky also shows how the conventional route of treating single diseases could, perversely, increase overall morbidity, because the longer people live, the greater the part played by ageing in health status. Interfering in mechanisms of ageing provides the current best prospect for preventing cancer, cardiovascular disease and dementia.

From a health-economics perspective, examined by Dana Goldman, the results of reducing morbidity in older people would depend on public policy. If older people were healthier and more active, then they would depend less on others and would produce greater economic activity — from work and volunteering — both of which would be a net benefit to society. A direct comparison of the predicted economic consequences of the status quo, delayed cancer, delayed heart disease and delayed ageing showed that this last scenario would result in a much higher proportion of able older people. However, for that to create an economic benefit, age-related entitlements to various forms of support, particularly pensions, would have to change.

Basic science and human demographic studies have delivered an unprecedented opportunity to tackle the comorbidities of later life. To translate these discoveries into drugs and changes in medical practice, and to reap the consequent economic benefits, will require some radical changes: breaking down disease siloes, training a new generation of physicians and scientists capable of working across disciplinary boundaries, and altering public attitudes and policy. ■

Linda Partridge is founding director of the Max Planck Institute for Biology of Ageing in Cologne, Germany, and director of the Institute of Healthy Ageing at University College London.
e-mail: linda.partridge@age.mpg.de



Aging: The Longevity Dividend

EDITED BY S. JAY OLSHANSKY, GEORGE M. MARTIN AND JAMES L. KIRKLAND
Cold Spring Harbor Laboratory Press: 2015.

Correspondence

Refugee fences fragment wildlife

Erecting border fences in parts of Europe in response to the current massive influx of refugees may harm wildlife. The fences can kill animals by entangling them in razor wire and will jeopardize the hard-won connectivity of species populations.

The human toll of the refugee crisis deserves the highest political attention. At the same time, many of the fences could be in violation of commitments under international conservation agreements, such as the European Commission's Habitats Directive.

With the opening of political borders during the twentieth century, Europe's large fauna have rebounded. This success is a result of trans-boundary conservation projects backed by legislation and effective management.

However, refugee fences have proliferated along the borders of Slovenia, Croatia and Hungary, for example, and more are planned along the boundaries of Latvia and Estonia with Russia. These are likely to affect brown bear, wolf, lynx and red deer species.

Mitigation measures should include adapting national conservation-management schemes to ensure the survival of newly isolated animal populations; designing the structure and placement of fences to minimize their impact on wildlife; and removing the fences at the earliest opportunity.

John D. C. Linnell* Norwegian Institute for Nature Research, Trondheim, Norway.
john.linnell@nina.no

*On behalf of 4 correspondents (see go.nature.com/fm6aaa for full list).

Treat wasting illness on multiple fronts

Cachexia is a complex wasting syndrome that cannot be fully reversed by nutritional support alone (see, for example, *Nature*

528, 182–183; 2015). There is accumulating evidence that a comprehensive multimodal approach may succeed where unimodal treatments (such as nutrition or anabolic drugs) have failed to deliver extended clinical benefits. Support for a multimodal policy comes from established rehabilitation programmes (see, for example, M. A. Spruit *et al. Am. J. Respir. Crit. Care Med.* 188, e13–e64; 2013).

The failure of the classical unimodal approach suggests that a shift in clinical-trial design is needed (K. C. H. Fearon *et al. J. Cachexia Sarcopenia Muscle* 6, 272–274; 2015). This could include monitoring the combined effects of exercise and nutrition, along with controlling metabolism and systemic inflammation. These interventions would need to be tested early, before cachexia becomes irreversible.

The complexity of such interventions makes them difficult to organize and fund. They would require input from research, government, pharmaceutical companies and regulatory authorities. However, the possible clinical benefits stand to improve the quality and, in the long term, perhaps even the quantity of patients' lives.

Kenneth Fearon University of Edinburgh, UK.
k.fearon@ed.ac.uk
Competing financial interests declared; see go.nature.com/npxkyg.

EU conservation overlooks geology

The European Commission needs to expand its conservation policy to protect its seriously threatened geological heritage.

Legislation for nature conservation in the European Union has so far focused mainly on biodiversity and habitats (see, for example, V. Hermoso *Nature* 528, 193; 2015). But fossils, rocks, minerals and

landforms also contribute to a country's geological landscape and heritage. Their features are a scientific asset that is shared by all countries, as well as an educational and cultural resource. They are also essential for supporting services to biodiversity. For example, geological sites in coastal cliffs and rocky outposts harbour and protect huge varieties of sea birds.

Neither of the two EU conservation directives that are currently under review (see go.nature.com/vkm9r7) includes the non-living elements of natural heritage, making it hard to encourage public respect for important geological features.

José Brilha European Association for the Conservation of Geological Heritage (ProGEO); and University of Minho, Braga, Portugal.
jbrilha@dct.uminho.pt

Plans for European medical doctorate

I agree with Stefan Hardt and colleagues on the benefits of a unified European medical doctorate (*Nature* 528, 333; 2015). However, removing the research component to create a vocational degree could result in a shortage of clinician scientists.

This is evident from our (unpublished) 2014 survey of 1,069 supervisors of dissertations at Charité in Berlin, one of Europe's largest university hospitals. Just under 1% of 3,714 research projects were of an MD-with-PhD type, more than two-thirds were MD projects and the rest were straight PhDs and dental or nursing projects. Thus, shifting to a purely vocational European medical doctorate system would mean many physicians missing out on useful research training (see also D. M. Milewicz *et al. J. Clin. Invest.* 125, 3742–3747; 2015).

One solution might be to

integrate the European medical doctorate into the medical curriculum. This would also encourage more-consistent application of evidence-based medicine in daily practice throughout the European Union (see J. Hilgers *et al. Med. Teach.* 29, 270–275; 2007). Formal training of lecturers, tutors and supervisors responsible for this integration would help to standardize and improve the quality of dissertation supervision (see *Nature* 527, 7; 2015).

Marc Dewey Charité — University Medicine Berlin, Germany.
dewey@charite.de

Monitor safety of aged fuel pipelines

Ensuring the integrity and safety of old pipelines that transport oil and natural gas calls for frequent inspections, together with modern, sensitive leak-detection tools and regular removal of accumulated deposits.

The failure of old pipelines is becoming increasingly common, and can be dangerously disruptive to communities and the environment. Examples from the United States include the rupture in 2010 of a 41-year-old oil pipeline in Michigan, which spilled around 4.5 million litres of oil into the Kalamazoo River, and the 2013 failure of a 65-year-old pipeline in Arkansas, requiring 22 homes to be evacuated.

Modern pipelines built from high-quality steels are statistically safer than transporting such fuels by road or rail (J. Behar and S. Al-Azem *World Pipelines* 15 (4), 18–28; 2015). However, over half of US underground pipelines are more than 50 years old (see go.nature.com/gczd6b). Such aged pipelines could fail at any time from corrosion, cracking or coating deterioration (X. Li *et al. Nature* 527, 441–442; 2015).

Frank Cheng University of Calgary, Alberta, Canada.
fcheng@ucalgary.ca

MICROBIOME

Fibre for the future

A chronic lack of dietary fibre has been found to reduce the diversity of bacteria in the guts of mice. This effect is not fully reversed when fibre is reintroduced, and increases in severity over multiple generations. [SEE LETTER P.212](#)

ERIC C. MARTENS

People living in industrialized nations routinely consume much less than the recommended amount of 25–38 grams of dietary fibre per day. Physicians and nutritionists have been imploring us for decades to bolster our fibre intake to help stave off maladies ranging from heart disease to intestinal disorders. The mechanisms through which fibre consumption modulates health are manifold, including a role in maintaining our resident gut microorganisms. On page 212 of this issue, Sonnenburg *et al.*¹ reveal that a lack of dietary fibre leads to a substantial loss of diversity in this microbial community, and influences the ability of gut bacteria to be transferred from parents to their offspring. Furthermore, it seems that simply restoring fibre consumption is not enough to reverse this effect once it has been passed to subsequent generations.

The 'fibre' that we see quantified on food labels is a catch-all category encompassing dozens of different molecules, mostly complex carbohydrates (linear and branched chains of simple sugars such as glucose). But the human genome encodes only around a dozen digestive enzymes that target complex carbohydrates. Technically speaking, dietary fibre comprises

the polymeric molecules that cannot be broken down by these enzymes. However, these nutrients do not go to waste. Instead, the diverse microorganisms that have evolved to inhabit the human intestine — collectively called the gut microbiota — produce thousands of enzymes that specifically target dietary fibre^{2–4}. Some individual bacteria produce more than 300 such enzymes⁵. These organisms ferment the released sugars into short-chain fatty acids, which are used as fuel for intestinal cells and which influence systemic physiology and the development of immune responses⁶.

The gut microbiota of each person typically contains hundreds of different bacterial species. We do not each harbour exactly the same community members; rather, the composition of our microbiota is drawn from a larger set of potential colonizers on the basis of parental and environmental exposure that begins at birth. Many microorganisms that live in the human gut exist only in this niche, and thus rely on successful transfer between generations to avoid extinction.

Sonnenburg *et al.* posed the question: what happens to the microbiota when dietary fibre is withheld for prolonged periods? The researchers colonized the intestines of germ-free mice (those that lack any resident microorganisms) with a human faecal sample, which contains a

representative complement of the gut microbiota members. They then fed the mice a diet rich in dietary fibre or one that contained only low fibre, in a form poorly accessible to the microbiota. After several weeks of fibre deprivation, the microbiota showed a reduction in the abundance of many bacterial groups that had been previously present (Fig. 1). These bacteria continued to thrive in the mice that were fed a high-fibre diet. When the fibre-starved mice were returned to a normal diet and allowed to recover for several weeks, many of these groups came back, but some failed to return to their previous levels, revealing that prolonged diet shifts can inflict changes that persist after dietary intervention.

The authors next investigated how fibre consumption affects the microbiota over multiple generations. They allowed the mice colonized with human bacteria, from both the high- and low-fibre cohorts, to breed within their cohorts, and for natural microbial colonization of the offspring to occur through maternal contact. Offspring born to parents fed the low-fibre diet had reduced microbiota diversity irrespective of whether they were weaned onto the same diet as their parents or onto a high-fibre diet. Strikingly, the reduction in gut bacterial diversity that was observed in the first generation was compounded over each

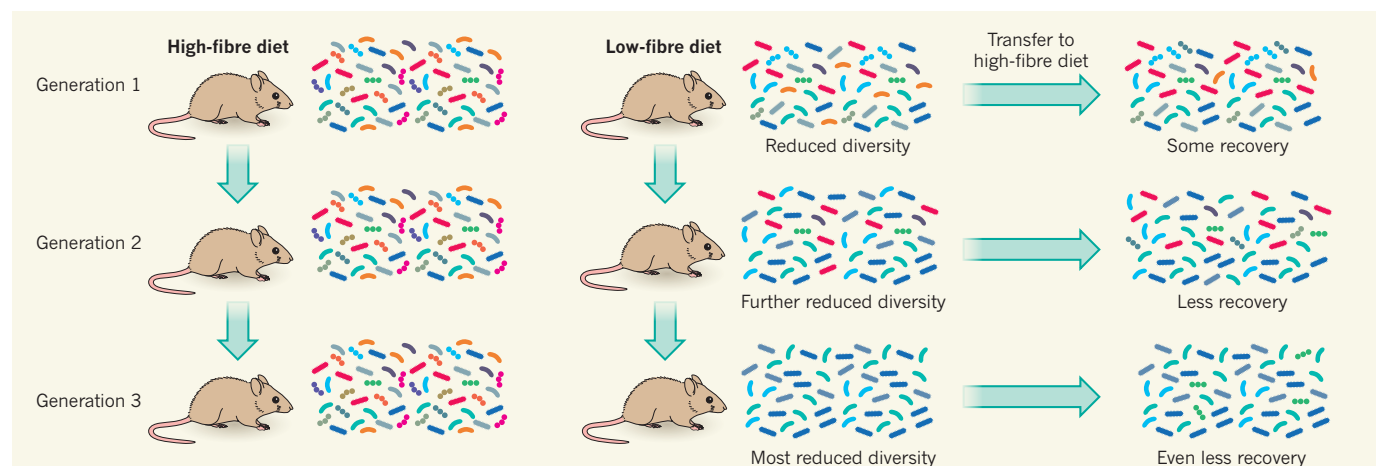


Figure 1 | Loss of diversity. Sonnenburg *et al.*¹ found that mice fed a low-fibre diet had a lower species diversity in their gut microbiota than mice fed a high-fibre diet. In first-generation mice, most (but not all) of this diversity was recoverable when mice on the low-fibre diet were switched to a high-fibre diet. However, the authors found that diversity loss was greater in each subsequent generation maintained on a low-fibre diet, and that the degree of recovery also decreased, implying extinction of some microbial species.

of four subsequent generations. Moreover, the inferred genomic content of the bacteria that remained after four generations suggested that the abundance of several fibre-degrading enzyme families had been reduced. But further work is required to find out whether a loss in fibre-degrading capacity occurred.

To assess whether dietary change might ameliorate these deficiencies, Sonnenburg *et al.* placed some of the mice from each generation of fibre-deprived mice on a high-fibre diet. The inability to recover lost diversity was a consistent characteristic at each generation (Fig. 1). However, transplanting the fibre-starved mice with a faecal sample from mice fed a high-fibre diet successfully restored most of the missing bacteria.

It is becoming increasingly apparent that the gut microbiota of people in cultures that eat less-processed and higher-fibre diets differ from those of people in industrialized countries, and often contain a higher diversity of microorganisms^{7–9}. Humans have co-evolved with symbiotic bacteria, and these microbial partners shoulder most of the burden of digesting complex carbohydrates. It remains to be determined whether some of this functionality has already been lost in some people and, if so, to what extent. However, in the future, we may turn to probiotic formulations, possibly derived from humans or animals that have not yet restricted their gut microbiome through a low-fibre diet, to restore essential functions that have been lost.

Carbohydrates frequently get a bad rap in fad diets, largely owing to simple carbohydrates such as glucose and fructose that permeate Western diets and provide us with an excess of easy calories. However, their complex cousins that are naturally present in plants, whole grains and a variety of other sources are worth consuming in greater amounts. Two authors of this study last year published a book for the popular press, *The Good Gut*¹⁰, which chronicles the interaction between diet, the microbiome and health, and is replete with high-fibre recipes. You just might consider choosing a salad at lunch today or an extra serving of beans at dinner. Future generations may thank you, too. ■

Eric C. Martens is in the Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA.
e-mail: emartens@umich.edu

1. Sonnenburg, E. D. *et al.* *Nature* **529**, 212–215 (2016).
2. Martens, E. C. *et al.* *PLoS Biol.* **9**, e1001221 (2011).
3. Larsbrink, J. *et al.* *Nature* **506**, 498–502 (2014).
4. Cuskin, F. *et al.* *Nature* **517**, 165–169 (2015).
5. El Kaoutari, A. *et al.* *Nature Rev. Microbiol.* **11**, 497–504 (2013).
6. Smith, P. M. *et al.* *Science* **341**, 569–573 (2013).
7. De Filippo, C. *et al.* *Proc. Natl Acad. Sci. USA* **107**, 14691–14696 (2010).
8. Clemente, J. C. *et al.* *Sci. Adv.* **1**, e1500183 (2015).
9. Schnorr, S. L. *et al.* *Nature Commun.* **5**, 3654 (2014).
10. Sonnenburg, E. & Sonnenburg, J. *The Good Gut* (Penguin, 2015).

COSMOLOGY

Photons from dwarf galaxy zap hydrogen

The detection of photons sufficiently energetic to ionize neutral hydrogen, coming from a compact, star-forming galaxy, offers clues to how the first generation of galaxies may have reionized hydrogen gas in the early Universe. [SEE LETTER P.178](#)

DAWN K. ERB

Most of the ordinary matter in the Universe is found not in stars, but in the diffuse gas between galaxies: the intergalactic medium (IGM), which is mainly hydrogen. This gas is almost completely ionized, and has been so since the formation of the first stars and galaxies a few hundred million years after the Big Bang. But few details are known about the sources of the radiation that ionized the gas, or how this radiation escaped from its source galaxies. On page 178 of this issue, Izotov *et al.*¹ report the detection of ionizing radiation from a star-forming dwarf galaxy in the local Universe, which may clarify the escape question.

The hydrogen gas that pervades the Universe has undergone phase changes over 13.8 billion years. The early Universe was too hot for protons and electrons to combine into neutral

hydrogen, so the hydrogen was ionized. The Universe cooled as it expanded, and about 375,000 years after the Big Bang, the temperature decreased enough for neutral hydrogen to form.

The gas remained neutral for the next few hundred million years, until the epoch of reionization — the last major phase transition in the Universe (Fig. 1). This transition occurred when the first sources of photons that were energetic enough to ionize hydrogen appeared in sufficient numbers to reionize the IGM. These photons are known as Lyman continuum photons because they have wavelengths shorter than the Lyman limit of 912 ångströms, which corresponds to the energy required to ionize the hydrogen atom (13.6 electronvolts).

We now know, from the scattering of cosmic microwave background photons by reionized electrons² and from observations of the

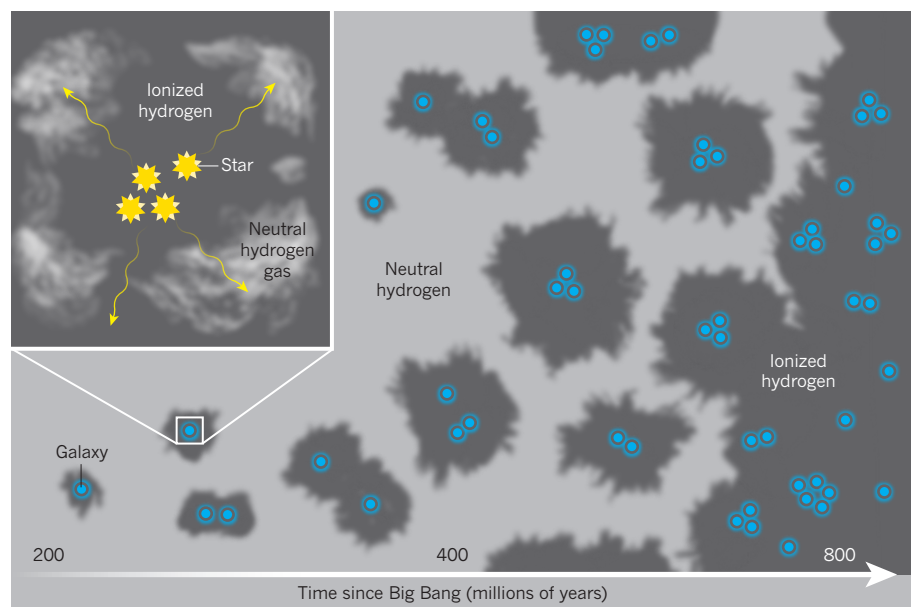


Figure 1 | Escape of ionizing radiation from a galaxy. The hydrogen gas between galaxies was ionized by the first stars and galaxies when the Universe was about 400 million years old. This probably occurred gradually (main panel): radiation from stars and galaxies ionized increasingly large bubbles of gas, until the ionized regions completely overlapped. For this to happen, photons with enough energy to ionize hydrogen must escape from galaxies. Such escape is challenging, because galaxies are filled with neutral hydrogen gas that absorbs ionizing radiation (inset; arrows represent ionizing photons). Izotov *et al.*¹ report the detection of ionizing photons from a compact starburst galaxy in the nearby Universe, a discovery that helps to explain the conditions that allow such radiation to escape from galaxies.

absorption by neutral hydrogen in the spectra of extremely distant quasars³, that reionization was a gradual process, the midpoint of which occurred approximately 400 million years after the Big Bang. However, we have not yet observed the changing ionization state of the IGM directly, and theoretical models struggle to explain how the known population of galaxies at this epoch could have produced enough radiation for reionization to occur.

The problem is twofold: large numbers of faint galaxies seem to be required to supply the necessary radiation⁴, but photons must also be able to escape from the galaxies in which they are produced. Stars form out of cool gas, and thus star-forming galaxies are filled with neutral hydrogen. Neutral hydrogen absorbs ionizing photons and therefore, even in a galaxy whose stars produce copious amounts of energetic radiation, the proportion of this radiation that actually escapes from the galaxy (the escape fraction) may be extremely low.

The Lyman continuum radiation from the galaxies that reionized the Universe will never reach our telescopes, because the photons are absorbed by the IGM long before they get to Earth. It is possible to detect ionizing radiation from closer galaxies, those corresponding to an age of the Universe of approximately 1.6 billion to 2 billion years (see refs 5 and 6, for example), but such detections are complicated by the probability of contamination by non-ionizing radiation originating in faint galaxies along the line of sight. In some ways, galaxies in the local Universe offer the best prospect for a detailed understanding of how ionizing radiation escapes from galaxies. Complicating factors are the unknown similarities and differences between local galaxies and the galaxies of the reionization era, and the fact that Earth's atmosphere blocks the ultraviolet wavelengths of this radiation, requiring that observations be made from space.

Izotov *et al.* focused on the compact dwarf galaxy J0925 + 1403 — deemed likely to produce escaping ionizing radiation on the basis of properties inferred from its optical emission lines. These lines indicate an unusually high ionization state in the gas near the galaxy's star-forming regions, suggesting that the stars may make more ionizing radiation than can be absorbed by the surrounding gas. The authors' successful detection of ionizing photons is the fourth such observation from a nearby galaxy^{7–9}.

Crucially, this galaxy has the highest escape fraction yet measured locally: about 8%, compared with the roughly 1–3% escape fraction measured from other nearby galaxies^{7–9}. The total amount of escaped radiation is sufficient to ionize a mass of IGM gas 40 times greater than the galaxy's stellar mass. Finding Lyman continuum radiation from this galaxy therefore broadly confirms our understanding of the general conditions that may facilitate the escape of ionizing radiation.

However, much work remains to be done to understand how galaxies reionized the Universe: the current study involves a single galaxy, whereas reionization depends on the properties of a population. It is not yet clear whether or not J0925 + 1403 is typical of compact, highly ionized starbursts (galaxies with extremely high rates of star formation) in the nearby Universe. We also do not know whether this galaxy is similar to those that reionized the Universe; its small size, high ionization state and relatively low degree of enrichment by elements heavier than helium generally match the expected properties of such objects, but none of these properties has been measured for the earliest galaxies.

Izotov *et al.* report that J0925 + 1403 leaks a large number of ionizing photons relative to its ultraviolet luminosity. This finding will inform future theoretical models of the reionization of the Universe by faint galaxies, but it remains to be determined whether this result is typical in the local Universe or representative of galaxies in the reionization era. The authors'

detection therefore emphasizes the need for additional, larger studies to develop a statistical understanding of Lyman continuum escape in the local Universe and its relationship to the properties of galaxies more generally. ■

Dawn K. Erb is at the Leonard E. Parker Center for Gravitation, Cosmology and Astrophysics, Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53211, USA. e-mail: erbd@uwm.edu

1. Izotov, Y. I. *et al.* *Nature* **529**, 178–180 (2016).
2. Planck Collaboration. *Astron. Astrophys.* **571**, A16 (2014).
3. Fan, X., Carilli, C. L. & Keating, B. *Annu. Rev. Astron. Astrophys.* **44**, 415–462 (2006).
4. Robertson, B. E. *et al.* *Astrophys. J.* **768**, 71 (2013).
5. Steidel, C. S., Pettini, M. & Adelberger, K. L. *Astrophys. J.* **546**, 665–671 (2001).
6. Vanzella, E. *et al.* *Astrophys. J.* **751**, 70 (2012).
7. Bergvall, N. *et al.* *Astron. Astrophys.* **448**, 513–524 (2006).
8. Leitert, E., Bergvall, N., Hayes, M., Linné, S. & Zackrisson, E. *Astron. Astrophys.* **553**, A106 (2013).
9. Borthakur, S., Heckman, T. M., Leitherer, C. & Overzier, R. A. *Science* **346**, 216–219 (2014).

VASCULAR BIOLOGY

Transcriptional control of endothelial energy

The formation of blood vessels requires rapid proliferation of endothelial cells. The transcription factors FOXO1 and MYC have been found to regulate the metabolism and proliferation of vascular endothelial cells. SEE LETTER P.216

CHRISTER BETSHOLTZ

Research during the past decade has yielded extensive knowledge of the cellular and molecular mechanisms of angiogenesis, the process through which new blood vessels form from existing ones. Although we have learnt much about how vessels sprout, elongate, branch, form lumens and regress¹, one piece of the angiogenic puzzle has remained poorly explored: how the proliferation of the endothelial cells that line the vessels' interior is regulated. In this issue, Wilhelm *et al.*² (page 216) show that the transcription factor FOXO1 couples growth-factor signalling to the metabolism, growth and division of endothelial cells. The authors also identify the protein MYC, a known driver of cancer development and the anabolic metabolism that constructs tissues³, as a key mediator of endothelial FOXO1 function.

Unicellular organisms have evolved to grow and divide whenever nutrients are abundant, and, conversely, to enter states of inactivity when nutrients are scarce. Multicellular organisms function differently. In our bodies, most

cells are exposed to a constant, rich supply of nutrients, but proliferate only when stimulated by growth factors, such as during organ development and regeneration. For endothelial cells, the situation differs again — these cells proliferate and form new blood vessels when oxygen and nutrients are low, with the aim of increasing oxygen and nutrient delivery to other cells in the tissues. Once functional vessels have been established, the endothelial cells, now also exposed to high levels of oxygen and nutrients, cease to proliferate.

Signalling induced when vascular endothelial growth factor A (VEGFA) binds to VEGF receptor 2 (VEGFR2) is the principal driver of most of the fundamental morphogenetic events involved in angiogenesis, including endothelial cell proliferation. A key pathway downstream of VEGFR2 is the PI3K–AKT pathway⁴ — a powerful regulator of glucose metabolism and protein synthesis⁵. The protein AKT also inhibits the activity of FOXO transcription factors by phosphorylating them: this causes them to be redistributed from the cell nucleus to the cytoplasm⁶ (Fig. 1).

Wilhelm *et al.* set out to test the idea that

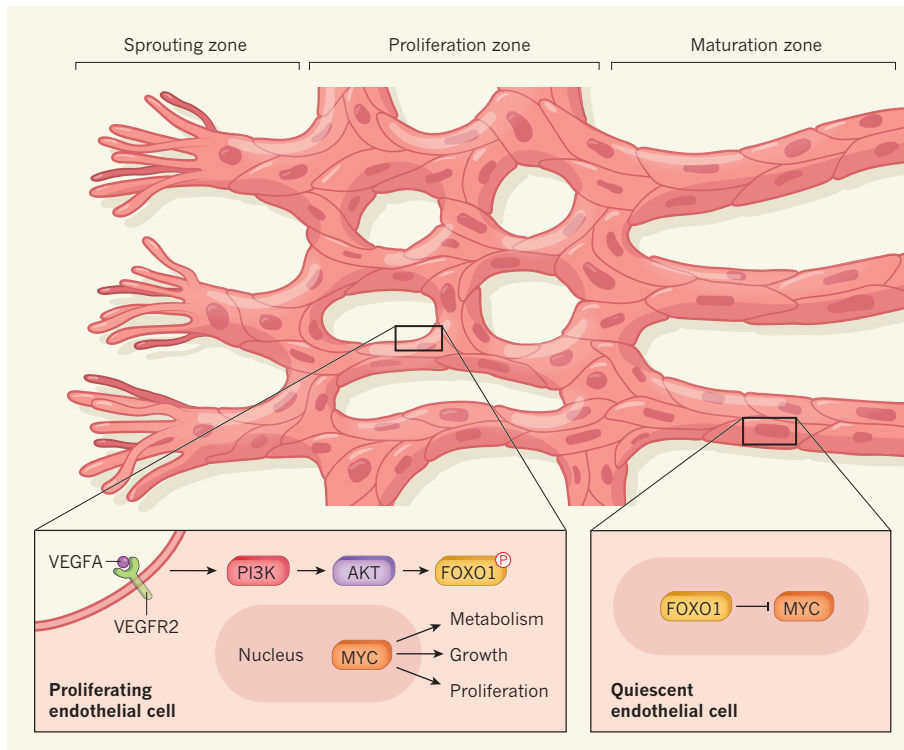


Figure 1 | FOXO1 and MYC contribute to regulating angiogenesis. As new blood vessels form, extensive proliferation of the endothelial cells that line them takes place just behind the sprouting tip of the vessel. Wilhelm *et al.*² show that, in these endothelial cells, the transcription factor FOXO1 is located in the cytoplasm, possibly as a result of having been phosphorylated (P) through the activity of the PI3K–AKT signalling pathway that is induced when vascular endothelial growth factor A (VEGFA) binds to its receptor (VEGFR2). As a consequence, FOXO1 cannot exert its function, described by the authors, in inhibiting the transcription factor MYC, which remains in the nucleus. The resulting enhanced MYC activity leads to increased cellular metabolism, growth and proliferation. By contrast, vessel maturation, which occurs more centrally in the vascular network, coincides with cessation of growth-factor signalling. Presumably at this stage, FOXO1, now non-phosphorylated, moves to the nucleus and inhibits MYC, thereby inducing endothelial-cell quiescence.

FOXO1, the FOXO family member enriched in endothelial cells, might constitute a link between growth-factor signalling input, metabolism and cell proliferation (Fig. 1). The researchers inactivated the *Foxo1* gene specifically in endothelial cells of newborn mice and found that this led to overgrowth of these cells and the formation of a hugely disorganized and dilated vascular network in the developing mouse retina. Conversely, endothelial-specific expression of a constitutively active FOXO1 protein resulted in a sparse retinal vasculature composed of fewer than normal endothelial cells.

Expression of constitutively active FOXO1 also led to decreased glucose uptake, glycolysis and lactate production in endothelial cells. The authors further observed decreases in oxygen consumption, the production of reactive oxygen species and levels of the energy-carrying molecule ATP — all features that correspond with reduced cellular metabolic activity. The cells survived, but entered a state of metabolic quiescence (a form of dormancy), accompanied by lower expression of genes that are targeted by the transcription factor MYC.

Because MYC is known to regulate all of the

above-mentioned metabolic processes, and because inhibition of MYC by FOXO occurs in other cells⁶, the authors tested whether MYC might be the mediator of the proliferation-stimulating effect of *Foxo1* inactivation. They found that constitutively active FOXO1 suppressed MYC expression and inactivation of *Foxo1* had the opposite effect. FOXO1 also increased expression of negative regulators of MYC activity, including MXI1 and FBXW7, suggesting that FOXO1 inhibits MYC at several levels. Furthermore, the authors show that MYC overexpression restored metabolism and proliferation in endothelial cells with constitutively active FOXO, and repaired vascular defects induced by this treatment. Together, these data provide compelling evidence for MYC as an effector of *Foxo1* deficiency in endothelial cells.

Wilhelm and colleagues' work reveals a central mechanism whereby the control of endothelial-cell proliferation is linked to the cells' metabolic state. But as with all good studies, it generates many questions. Does the amount of FOXO1 change during angiogenesis, or is its activity regulated solely by transport in and out of the nucleus, as the

authors' results might suggest (see Figure 1a of the paper²)? Although the experimental methods used by the authors are the best available, they involved vast changes in FOXO1 levels (complete loss or several-fold increase), which would not occur in physiological settings. The striking normalization of the retinal vasculature observed in mice overexpressing both constitutively active FOXO1 and MYC might result from a new balance, achieved through similarly increased levels of the two. A more critical test of the role of FOXO1 as a 'rheostat' of vascular expansion, a term used by the authors, should ideally include manipulations of the activity of FOXO1 — such as of its nuclear–cytoplasmic shuttling — at normal levels.

Moreover, the question of how FOXO1 nuclear translocation is regulated in endothelial cells remains unresolved. In analogy with the regulation of other FOXO proteins by a growth-factor–PI3K–AKT axis, one would guess that signalling downstream of VEGFA–VEGFR2 binding plays a central part. However, FOXOs have several upstream inputs besides AKT⁶, and FOXO1 signalling and its role in endothelial cells might be multifaceted and context-dependent. For example, in adult mice, deletion of multiple FOXO proteins, including FOXO1, leads to the formation of benign endothelial-cell tumours known as haemangiomas in some organs, but not all⁷. Similarly, VEGF induces enhanced proliferation in cultures of non-FOXO-expressing endothelial cells from some organs, but not others⁷.

Although this association between FOXO dysregulation and endothelial tumour formation concurs with Wilhelm and colleagues' idea that FOXO1 is a major regulator of endothelial-cell proliferation, and extend their observations into adult animals, the non-uniform and tissue-type-dependent responses are intriguing. Further study of FOXOs is warranted, particularly in the regulation of endothelial metabolism and proliferation at different stages of development and different vascular sites. ■

Christer Betsholtz is in the Department of Immunology, Genetics and Pathology at Uppsala University, and the Department of Medical Biochemistry and Biophysics at the Karolinska Institutet, Stockholm, Sweden. e-mail: christer.betsholtz@igp.uu.se

1. Potente, M., Gerhardt, H. & Carmeliet, P. *Cell* **146**, 873–887 (2011).
2. Wilhelm, K. *et al. Nature* **529**, 216–220 (2016).
3. Stine, Z. E., Walton, Z. E., Altman, B. J., Hsieh, A. L. & Dang, C. V. *Cancer Discov.* **5**, 1024–1039 (2015).
4. Olsson, A.-K., Dimberg, J., Kreuger, J. & Claesson-Welsh, L. *Nature Rev. Mol. Cell Biol.* **7**, 359–371 (2006).
5. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. *Science* **324**, 1029–1033 (2009).
6. Eijkelboom, A. & Burgering, B. M. T. *Nature Rev. Mol. Cell Biol.* **14**, 83–97 (2013).
7. Paik, J.-H. *et al. Cell* **128**, 309–323 (2007).

This article was published online on 6 January 2016.

CLIMATE SCIENCE

Earth's narrow escape from a big freeze

An equation has been derived that allows the timing of the onset of glaciations to be predicted. This confirms that Earth has just missed entering a new glacial period, and is unlikely to enter one for another 50,000 years. [SEE LETTER P.200](#)

MICHEL CRUCIFIX

A subject of much debate is whether atmospheric levels of carbon dioxide were already significantly altered by emissions associated with human activities before the Industrial Revolution in the eighteenth century. One estimate suggests that the atmospheric concentration of CO_2 would have been only 240 parts per million (p.p.m.) in an agriculture-free world, rather than 280 p.p.m., as was measured just before the Industrial Revolution¹. On page 200 of this issue, Ganopolski *et al.*² report modelling studies confirming that we would now be entering an ice age if the concentration had remained at 240 p.p.m. By contrast, they report that glacial inception — the onset of an ice age — could not have occurred at CO_2 concentrations that were typical of the eighteenth century.

The Quaternary period has conventionally been divided into two epochs: the Pleistocene, which lasted from about 2.59 million to 12,000 years ago, and the Holocene, which followed the Pleistocene and continues to the present day. The Pleistocene was a time of great, successive glaciations interspersed with interglacial periods, during which environmental conditions were similar to those occurring today. During the Holocene — the latest interglacial period — humans invented agriculture, and their impact on the environment increased at an exponential rate. One of the signatures of this impact is the rising concentration of CO_2 in the atmosphere. But at what point does this impact become sufficiently large to affect climate and glacial inception?

A modelling study³ in 2000 established that pre-industrial levels of CO_2 were high enough to guarantee a period of interglacial

conditions for at least 50,000 years (Fig. 1). Consistent with Ganopolski and colleagues' findings, this earlier study also predicted that the next glacial inception (which would have led to a glaciation that reached an ice maximum 60,000 years from now), could not now occur owing to the warming effect of anthropogenic emissions. Moreover, probabilistic assessments^{4,5} of the timing of the next glacial inception have been provided by

using simple dynamical systems for climate prediction, calibrated using data about past CO_2 levels and ice volumes. All of these studies used different models and assumptions, but they broadly agree on the potential timing for a glacial inception because their forecasts are determined by predictable drops in incoming solar radiation (insolation) in the Northern Hemisphere caused by changes in Earth's orbit.

Ganopolski and co-workers' study is an advance on previous work because it provides a simple equation for predicting when glacial inception will occur. The researchers observed that, in the Earth-system model they used for their study (CLIMBER-2), ice begins to form when insolation in the Northern Hemisphere at the summer solstice falls below a certain value that depends logarithmically on the concentration of atmospheric CO_2 . They were thus able to work out an equation that describes this behaviour.

To calibrate the equation, the authors performed several simulations that differed by the value of a parameter that controls cloud height in their model. This sampling process effectively generates a family of model versions, which the authors tested to see which ones predicted past glacial inceptions. Past glaciations and interglacials have been identified on the basis of isotopic data from marine sediments, and they follow a numbering scheme in which isotope 'stages' with odd numbers roughly correspond to interglacials. The authors paid special attention to the glacial inceptions after marine isotope stages 19 and 11, and to the period after marine isotope stage 1 (that is, the Holocene), because insolation evolved in a similar way at those times but led to different outcomes (stage 1 did not produce a glacial inception). Only the parameter values that yielded correct simulations of all past glacial inceptions were used to establish the equation.

The authors were thus able to confirm that Earth had a narrow escape from glacial inception during the Holocene: the increase in atmospheric CO_2 levels during this period was sufficient to prevent the planet from entering a glacial period. The authors also report that an interglacial climate would have continued for at least 20,000 years, and more plausibly for 50,000 years, if CO_2 concentrations had been sustained at levels typical of the eighteenth century. However, almost 500 gigatonnes of carbon (GTC; 1 GTC is equivalent

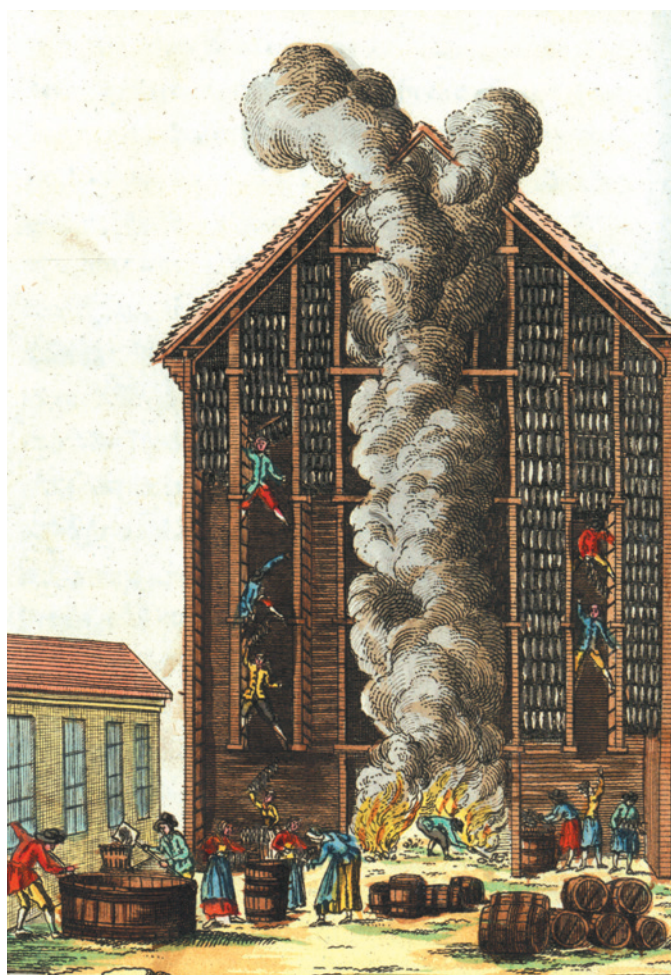


Figure 1 | An eighteenth-century smokehouse. The atmospheric level of carbon dioxide just before the Industrial Revolution was 280 parts per million and may already have been affected by emissions associated with human activities. Ganopolski *et al.*² report models suggesting that atmospheric CO_2 levels typical of the eighteenth century were high enough to prevent the onset of a glacial period for 50,000 years.

FLORILEGIUS/ALAMY

to 3.6 gigatonnes of CO₂) have been released into the atmosphere since the Industrial Revolution. Ganopolski *et al.* show that this means that we will probably skip the next glacial inception too: emissions of 1,000 GTC (a scenario that is quite likely) will almost guarantee 100,000 years without any glaciation.

Such long-term consequences may seem surprising, given that the emissions will occur over a few centuries at most and that anthropogenic CO₂ will eventually be absorbed by the oceans. But for this absorption to occur, carbonate minerals in the ocean will need to be dissolved, to counteract the increase in ocean acidity that occurs when CO₂ is absorbed, and which limits the amount of CO₂ that can be dissolved. This takes time. In fact, the mean half-life of CO₂ in the atmosphere is of the order of 35,000 years⁶. Consequently, anthropogenic CO₂ will still be in the atmosphere in 50,000 years' time, and even 100,000 years, which is enough to prevent any glaciation.

The method used by Ganopolski *et al.* is known as 'perturbed physics' sampling. This means that the different scenarios for future climate were sampled by modifying one parameter, which controls one of the physical effects described by the model. But no model is perfect, and all the possible errors associated with the model cannot be entirely compensated for by adjusting this parameter. To provide better predictions, we need to pay special attention to climate processes that are currently not well quantified.

Among them, the causes of CO₂ changes during past interglacial periods and during the early stages of glaciation remain a matter of controversy. For example, we are uncertain about the amplitude and dynamics of carbon sequestered in peatlands^{1,7}. More fundamentally, we do not yet know whether natural CO₂ dynamics have an active role in causing glacial inception, or whether they passively amplify the effects of accumulating ice at northern high latitudes. In spite of these uncertainties, Ganopolski and colleagues' main conclusion is likely to stand. It reinforces previous assessments asserting that humanity's collective footprint on Earth already extends beyond any imaginable future of our society. ■

Michel Crucifix is at the Université catholique de Louvain, Earth and Life Institute, Louvain-la-Neuve 1348, Belgium. e-mail: michel.crucifix@uclouvain.be

1. Ruddiman, W. F. *Annu. Rev. Earth Planet. Sci.* **41**, 45–68 (2013).
2. Ganopolski, A., Winkelmann, R. & Schellnhuber, H. J. *Nature* **529**, 200–203 (2016).
3. Loutre, M. F. & Berger, A. *Clim. Change* **46**, 61–90 (2000).
4. Hargreaves, J. C. & Annan, J. D. *Clim. Dynam.* **19**, 371–381 (2002).
5. Crucifix, M. & Rougier, J. *Eur. Phys. J. Spec. Top.* **174**, 11–31 (2009).
6. Archer, D. J. *Geophys. Res. Oceans* **110**, C09S05 (2005).
7. Menviel, L. & Joos, F. *Paleoceanography* **27**, PA1207 (2012).

ECOLOGY

A trail map for trait-based studies

Global assessments of variation in plant functional traits and the way that these traits influence competitive interactions provide a launching pad for future ecological studies. [SEE ARTICLE P.167](#) & [LETTER P.204](#)

JONATHAN M. LEVINE

Ecologists explore the processes that govern the natural world around us. But this can seem an uphill battle when nature presents such a wide diversity of species, each with its own set of interactions with the environment. One way to make sense of this diversity and its mechanistic underpinnings is to focus not on species but on the functional traits they possess, such as plant height, seed size or leaf area. Two papers in this issue advance our understanding of how traits vary between plant species, and the ramifications of this variation for competitive interactions. Díaz *et al.*¹ (page 167) document the patterns of functional-trait variation among plant species worldwide and reveal fundamental constraints on plant form that allow the organisms to survive natural selection, physiological challenges and competitive exclusion. Kunstler *et al.*² (page 204) show how functional traits consistently predict the competitive interactions between trees in six forested biomes, with effects counter to expectations from classic theory.

In the nineteenth century, Alexander von Humboldt and Charles Darwin wrote at length about the surprising diversity in form and function of organisms on Earth, and this topic still intrigues naturalists today. For much of the history of ecology, most patterns of diversity and abundance have been studied at the species level. But a growing number of ecologists, and plant ecologists in particular, think that studies focused on functional traits present greater opportunity for generality and predictability, and a tighter connection to organismal function^{3–6}. Indeed, it is a species' traits that determine the organism's growth, dynamics and interactions, not its taxonomic nomenclature. The implication is that a more productive way of asking, for example, the classic question of what processes maintain the diversity of species is to ask what processes explain the dispersion of traits among community members.

Díaz *et al.* have laid the groundwork for this approach and a wide range of ecological and evolutionary investigations by quantifying the dimensionality of plant 'trait space' — the multivariate space in which a plant can be

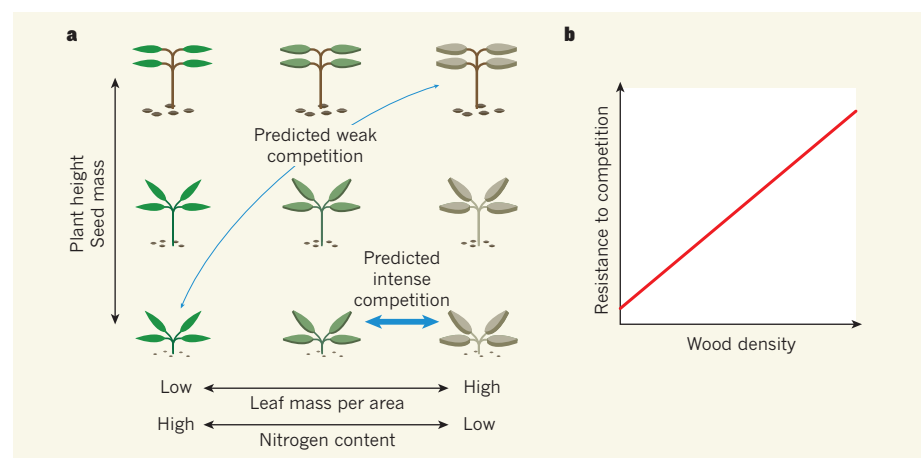


Figure 1 | Trait dimensions and competition. a, Díaz *et al.*¹ conducted an analysis of six plant traits across more than 45,000 species and found that most of the variation between species in this six-dimensional trait space lies along a two-dimensional plane. One dimension describes variation in leaves having an acquisitive strategy (low leaf mass per unit area and high nitrogen content) versus a conservative one (high leaf mass per unit area and low nitrogen content)⁶. The other dimension corresponds to variation in plant size (height and seed mass). Following classic ecological theory, species that are more similar to one another in the trait space should compete more intensely than would more different taxa. b, However, by working with a global data set of forest-tree growth, Kunstler *et al.*² found little to no support for this expectation. Instead, certain trait values were predictive of competitive superiority, such as greater wood density being associated with greater resistance to competition.

located on the basis of its functional traits. The researchers focused on six traits that are readily available for a large number of taxa and that are central to determining a plant's ecological strategy: plant height, leaf area, leaf mass per area, nitrogen content per mass, stem mass per volume and seed mass. Their analysis, which incorporates these values from more than 45,000 species, is the first of its scope to explore relationships across seed, leaf, stem and whole-plant traits. This scale was possible only because of a database called TRY (ref. 7), which contains 5.6 million records of plant functional traits assembled over the past decade.

In principle, a given plant could occupy any point within this six-trait space. To assess how constrained plant species actually are within this space, the authors compared their observations with four null models representing different distributions of and correlations between traits. They found that, worldwide, plant species occupy only a small fraction of their potential trait space, and the observed pattern is driven largely by strong correlations between functional traits across species (Fig. 1). The researchers then conducted a principal component analysis and identified two primary dimensions in which plants vary globally: plant size, ranging from short species with small seeds to tall species with large seeds; and leaf strategy⁶, ranging from 'acquisitive' species with low leaf mass per unit area and high leaf nitrogen content to 'conservative' species with high leaf mass per unit area and low nitrogen content.

It is difficult to evaluate the extent to which the constraints on plant form and function suggested by this analysis arise from biomechanical trade-offs, natural selection or competition. But this is where Kunstler and colleagues' study comes in. These researchers explored how three plant functional traits (leaf area per mass, plant height and wood density) predict the competitive interactions between forest tree species. Their data set is similarly impressive in scope to that of Díaz *et al.* — it includes three functional traits and trunk-diameter growth for more than 3 million trees from over 2,500 species in forest plots from 6 biomes. Taking advantage of natural variation in the density and identity of competitors surrounding a focal tree, the authors built a statistical model to quantify how a species' trait values predict its growth without competition, its resistance to competition and its ability to suppress the growth of neighbours.

The authors had good reason to expect that functional traits would predict competitive dynamics. Ecological theory holds that trait differences between species should cause these taxa to use the environment in different ways, resulting in a 'niche difference' that minimizes competition between species (Fig. 1a). Contrary to these expectations, however, Kunstler and colleagues found little to no evidence that trait differences minimize competition

between trees across the six forest biomes. Instead, certain trait values tended to predict the competitive advantage of one species over others. Trees with high wood density, for example, tended to be most resistant to competition (Fig. 1b). These findings resonate with work^{8,9} positing that trait differences should predict both the niche differences that stabilize species coexistence and the competitive imbalances that drive species exclusion.

If the three functional traits studied by Kunstler *et al.* predicted only competitive imbalances between forest trees, which traits explain their local coexistence? Although the authors point to trade-offs between high growth rate and competitive tolerance, their finding of greater competition within species than between species — a factor that also stabilizes species coexistence — must relate to traits other than those measured. This finding highlights the limitations of trait-based ecology. Although, in principle, all competitive dynamics must be explainable by plant traits, whether these are the functional traits that can be readily measured is an open question.

Answering this question, as well as related ones at the population and ecosystem levels, will require further integration of functional traits and mathematical models along the

lines of Kunstler and colleagues' approach. Nonetheless, ecologists have a growing need to efficiently predict the nature of competition between plant species that do not co-occur today but will in the future as climate change causes species to migrate, and to migrate at different rates¹⁰. Trait-based approaches may be our only recourse. ■

Jonathan M. Levine is at the Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zurich, 8092 Zurich, Switzerland.
e-mail: jlevine@ethz.ch

1. Díaz, S. *et al.* *Nature* **529**, 167–171 (2016).
2. Kunstler, G. *et al.* *Nature* **529**, 204–207 (2016).
3. McGill, B. J., Enquist, B. J., Weiher, E. & Westoby, M. *Trends Ecol. Evol.* **21**, 178–185 (2006).
4. Ackerly, D. D. & Cornwell, W. K. *Ecol. Lett.* **10**, 135–145 (2007).
5. Violle, C. *et al.* *Oikos* **116**, 882–892 (2007).
6. Wright, I. J. *et al.* *Nature* **428**, 821–827 (2004).
7. Kattge, J. *et al.* *Glob. Change Biol.* **17**, 2905–2935 (2011).
8. Adler, P. B., Fajardo, A., Kleinhesselink, A. R. & Kraft, N. J. B. *Ecol. Lett.* **16**, 1294–1306 (2013).
9. Kraft, N. J. B., Godoy, O. & Levine, J. M. *Proc. Natl Acad. Sci. USA* **112**, 797–802 (2015).
10. Alexander, J. M., Diez, J. M. & Levine, J. M. *Nature* **525**, 515–518 (2015).

This article was published online on 23 December 2015.

NEUROBIOLOGY

Pull out the stops for plasticity

The strength of synaptic connections between neurons needs to be variable, but not too much so. Evidence now indicates that regulation of such synaptic plasticity involves a complex cascade of feedback loops.

CHRISTINE E. GEE & THOMAS G. OERTNER

Learning is thought to manifest in the brain as physical changes that alter the strength of neuronal contact points called synapses. These contact points allow information to be transmitted from one neuron to another, and understanding the conditions that cause synapses to change strength (a phenomenon known as synaptic plasticity) has been a focus of neuroscience research for many years. Writing in *Nature Communications*, Tigaret *et al.*¹ challenge the prevailing idea that the local concentration of calcium ions (Ca^{2+}) is the key factor that determines whether a synapse becomes stronger or weaker after repetitive activation. They propose that plasticity involves an intracellular signalling cascade that overrides a safety mechanism. This suggests that the default state of the synapse is not to be plastic.

The main excitatory neurotransmitter in the mammalian brain is the molecule glutamate. Glutamate is released from the presynaptic neuron, and the postsynaptic neuron is excited when the molecule binds to and activates specialized receptor proteins, most of which are ion channels called ionotropic glutamate receptors. When activated, these channels open and positively charged ions enter the cell, depolarizing (reducing the voltage across) the cell membrane. In addition, glutamate receptors that are not ion channels, called metabotropic glutamate receptors, activate various intracellular signalling cascades. Their effect on synaptic transmission is generally slower than that of ionotropic receptors, but they are crucial for healthy brain function².

In the neuronal structure known as the dendritic spine, which forms a single synaptic contact, the initial depolarization caused by activation of ionotropic glutamate receptors

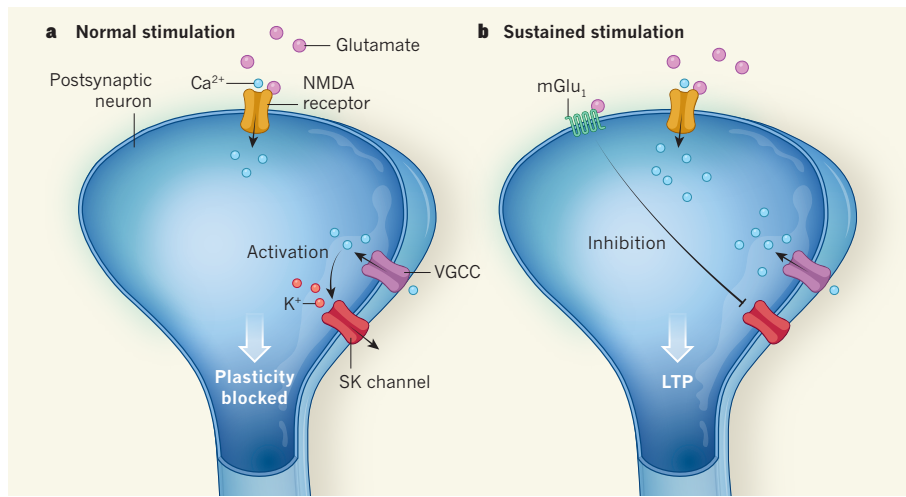


Figure 1 | The promotion of plasticity. **a**, The molecule glutamate is transmitted across the synaptic cleft between neurons to activate the postsynaptic neuron. When the voltage across the cell membrane decreases (depolarization), glutamate-bound NMDA-receptor proteins and voltage-gated calcium channels (VGCCs) open, allowing calcium ions (Ca²⁺) to enter the cell. Under normal conditions, proteins called SK channels are activated by this Ca²⁺ influx. Potassium ions (K⁺) flow out through SK channels, decreasing depolarization and preventing changes in synaptic strength, known as plasticity. **b**, Tigaret *et al.*¹ report that the metabotropic glutamate receptor protein mGlu₁ is activated by sustained glutamate signalling, and leads to inhibition of SK channels. This slow-acting inhibition enables prolonged depolarization and triggers strengthening of the synapse, known as long-term potentiation (LTP).

can be amplified by the opening of voltage-gated calcium channels, further depolarizing the spine. A special class of ionotropic glutamate receptors called NMDA receptors have a similar role — they open only when the neuron is already depolarized, forming a positive-feedback loop that increases Ca²⁺ influx and depolarization³. The activation of NMDA receptors is essential for many forms of long-lasting synaptic plasticity.

However, positive-feedback loops are inherently dangerous for neurons — too much depolarization and Ca²⁺ can be toxic, eventually triggering cell death. To prevent this from happening, spines have a safety mechanism in the form of calcium-activated, potassium-conducting SK channels⁴. When intracellular Ca²⁺ reaches a critical concentration, SK channels open, allowing positively charged potassium ions to exit the cell and so preventing further depolarization (Fig. 1a). This SK mechanism stops the positive-feedback loop, blocking further Ca²⁺ influx. But a side effect is that the synaptic strength becomes difficult to change⁵.

Tigaret *et al.* describe a plasticity-enabling mechanism that inhibits SK channels in individual spines. They found that repeated sequential activation of pre- and postsynaptic neurons in slices of rat brains induced synaptic strengthening, also known as long-term potentiation (LTP). In addition to NMDA-receptor activation, LTP induction required the activity of group 1 metabotropic glutamate receptors (mGlu₁). The authors show that activation of mGlu₁ triggers a slow-acting mechanism that inhibits SK channels, allowing for sustained depolarization and enhanced

Ca²⁺ entry into the spine (Fig. 1b).

The authors used a strong induction protocol (300 paired activations in 1 minute) to allow the relatively slow metabotropic process to take effect and enable LTP. This might seem unusual — after all, we don't need to be presented with information 300 times before learning a new association. Why was such a strong protocol required?

In an intact brain, specific neuromodulator chemicals such as dopamine and acetylcholine are released when the animal is in an aroused state: for example, when it learns that a certain sound predicts a frightening event. These substances modulate glutamate-activated synapses and have been shown to promote synaptic plasticity by blocking SK channels^{5,6}. The timing window for successful induction of LTP has been shown⁷ to change radically in the presence of neuromodulators. Thus, it seems that there is not just one rule for how synapses change during learning, but a whole set that are tailored to various occasions such as different mental states. This makes sense from a systems perspective — synaptic potentiation is gated not only by timing, but also by the brain's reward system. From an experimental point of view, the lack of neuromodulatory inputs, which is an inherent limitation of brain-slice experiments, might explain why a strong protocol was required. Functional imaging of single synapses in live, active animals is not yet possible.

The mechanism highlighted by Tigaret and colleagues is not the only way in which synapses can be strengthened. Enzymes called Src tyrosine kinases (which add phosphate groups to proteins) can directly enhance NMDA-receptor function⁸. This pathway has

been shown to cause LTP in the same type of synapse as that analysed in the current study⁹. The activity of various metabotropic receptors, including mGlu₁, can increase glutamate-mediated responses through this pathway¹⁰. It will be interesting to investigate whether the mGlu₁-triggered blockade of SK channels identified by Tigaret *et al.* acts together with direct NMDA-receptor phosphorylation to enable LTP, or whether one mechanism is dominant under specific conditions, depending, for instance, on cell type or the age of the animal.

This study also confirms¹¹ that, contrary to general thinking, it is not possible to predict the direction and magnitude of synaptic plasticity by simply analysing levels of Ca²⁺ in dendritic spines. For example, a pair of presynaptic stimulations triggered a very strong Ca²⁺ influx into the spine, but no plasticity whatsoever. But before Ca²⁺ is discarded as the key state variable, we must consider that successful induction of long-term plasticity relies on the interplay of local synaptic Ca²⁺ signals with Ca²⁺ signals in the cell body (soma) of the neuron. Indeed, the authors emphasize the importance of postsynaptic electrical activity and the activation of voltage-gated calcium channels for LTP. These processes are not restricted to the active spine; they increase Ca²⁺ levels throughout the neuron. Thus, it might be possible to predict the future strength of a synapse from simultaneous Ca²⁺ measurements in the spine and soma. This is certainly not an easy experiment, but sophisticated 3D scanning microscopes could be used to analyse compartmentalized Ca²⁺ signalling in individual neurons — and perhaps one day in intact animals during learning. ■

Christine E. Gee and Thomas G. Oertner
are at the Center for Molecular Neurobiology Hamburg (ZMNH), University Medical Center Hamburg-Eppendorf, 20251 Hamburg, Germany.
e-mail: thomas.oertner@zmnh.uni-hamburg.de

1. Tigaret, C. M., Olivio, V., Sadowski, J. H. L. P., Ashby, M. C. & Mellor, J. R. *Nature Commun.* <http://dx.doi.org/10.1038/ncomms10289> (2016).
2. Wang, H. & Zhuo, M. *Front. Pharmacol.* **3**, 189 (2012).
3. Bloodgood, B. L. & Sabatini, B. L. *Curr. Opin. Neurobiol.* **17**, 345–351 (2007).
4. Ngo-Anh, T. J., Bloodgood, B. L., Lin, M., Sabatini, B. L., Maylie, J. & Adelman, J. P. *Nature Neurosci.* **8**, 642–649 (2005).
5. Buchanan, K. A., Petrovic, M. M., Chamberlain, S. E. L., Marrior, N. V. & Mellor, J. R. *Neuron* **68**, 948–963 (2010).
6. Giessel, A. J. & Sabatini, B. L. *Neuron* **68**, 936–947 (2010).
7. Brzosko, Z., Schultz, W. & Paulsen, O. *eLife* **4**, e09685 (2015).
8. Yu, X. M., Askalan, R., Keil, G. J. & Salter, M. W. *Science* **275**, 674–678 (1997).
9. Lu, Y. M., Roder, J. C., Davidow, J. & Salter, M. W. *Science* **279**, 1363–1367 (1998).
10. Benquet, P., Gee, C. E. & Gerber, U. J. *Neurosci.* **22**, 9679–9686 (2002).
11. Nevian, T. & Sakmann, B. *J. Neurosci.* **26**, 11001–11013 (2006).

The global spectrum of plant form and function

Sandra Díaz¹, Jens Kattge^{2,3}, Johannes H. C. Cornelissen⁴, Ian J. Wright⁵, Sandra Lavorel⁶, Stéphane Dray⁷, Björn Reu^{8,9}, Michael Kleyer¹⁰, Christian Wirth^{2,3,11}, I. Colin Prentice^{5,12}, Eric Garnier¹³, Gerhard Bönsch², Mark Westoby⁵, Hendrik Poorter¹⁴, Peter B. Reich^{15,16}, Angela T. Moles¹⁷, John Dickie¹⁸, Andrew N. Gillison¹⁹, Amy E. Zanne^{20,21}, Jérôme Chave²², S. Joseph Wright²³, Serge N. Sheremet'ev²⁴, Hervé Jactel^{25,26}, Christopher Baraloto^{27,28}, Bruno Cerabolini²⁹, Simon Pierce³⁰, Bill Shipley³¹, Donald Kirkup³², Fernando Casanoves³³, Julia S. Joswig², Angela Günther², Valeria Falczuk¹, Nadja Rüger^{3,23}, Miguel D. Mahecha^{2,3} & Lucas D. Gorné¹

Earth is home to a remarkable diversity of plant forms and life histories, yet comparatively few essential trait combinations have proved evolutionarily viable in today's terrestrial biosphere. By analysing worldwide variation in six major traits critical to growth, survival and reproduction within the largest sample of vascular plant species ever compiled, we found that occupancy of six-dimensional trait space is strongly concentrated, indicating coordination and trade-offs. Three-quarters of trait variation is captured in a two-dimensional global spectrum of plant form and function. One major dimension within this plane reflects the size of whole plants and their parts; the other represents the leaf economics spectrum, which balances leaf construction costs against growth potential. The global plant trait spectrum provides a backdrop for elucidating constraints on evolution, for functionally qualifying species and ecosystems, and for improving models that predict future vegetation based on continuous variation in plant form and function.

Vascular plants are the main entry point for energy and matter into the Earth's terrestrial ecosystems. Their Darwinian struggle for growth, survival and reproduction in very different arenas has resulted in an extremely wide variety of form and function, both across and within habitats. Yet it has long been thought^{1–8} that there is a pattern to be found in this remarkable evolutionary radiation—that some trait constellations are viable and successful whereas others are not.

Empirical support for a strongly limited set of viable trait combinations has accumulated for traits associated with single plant organs, such as leaves^{7,9–12}, stems^{13,14} and seeds^{15–17}. Evidence across plant organs has been rarer, restricted geographically or taxonomically, and often contradictory^{18–29}. How tightly whole-plant form and function are restricted at the global scale remains unresolved.

Here we present the first global quantitative picture of essential functional diversity of extant vascular plants. We quantify the volume, shape and boundaries of this functional space via joint consideration of six traits that together capture the essence of plant form and function: adult plant height, stem specific density, leaf size expressed as

leaf area, leaf mass per area, leaf nitrogen content per unit mass, and diaspore mass. Our dataset, based on a recently updated communal plant trait database³⁰, covers 46,085 vascular plant species from 423 families and to our knowledge spans the widest range of growth-forms and geographical locations to date in published trait analyses, including some of the most extreme plant trait values ever measured in the field (Table 1, Extended Data Fig. 1). On this basis we reveal that the trait space actually occupied is strongly restricted as compared to four alternative null hypotheses. We demonstrate that plant species largely occupy a plane in the six-dimensional trait space. Two key trait dimensions within this plane are the size of whole plants and organs on the one hand, and the construction costs for photosynthetic leaf area, on the other. We subsequently show which sections of the plane are occupied, and how densely, by different growth-forms and major taxonomic groups. The design opportunities and limits indicated by today's global spectrum of plant form and function provide a foundation to achieve a better understanding of the evolutionary trajectory of vascular plants and help frame and test hypotheses as to where and

¹Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET and CEEFYn, Universidad Nacional de Córdoba, Casilla de Correo 495, 5000 Córdoba, Argentina. ²Max Planck Institute for Biogeochemistry, Hans-Knöll-Straße 10, 07745 Jena, Germany. ³German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany. ⁴Systems Ecology, Department of Ecological Science, Vrije Universiteit, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands. ⁵Department of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia. ⁶Laboratoire d'Ecologie Alpine, UMR 5553, CNRS – Université Grenoble Alpes, 38041 Grenoble Cedex 9, France. ⁷Laboratoire de Biométrie et Biologie Evolutive, UMR5558, Université Lyon 1, CNRS, F-69622 Villeurbanne, France. ⁸Institute of Biology, University of Leipzig, Johannisallee 21, 04103 Leipzig, Germany. ⁹Escuela de Biología, Universidad Industrial de Santander, Cra. 27 Calle 9, 680002 Bucaramanga, Colombia. ¹⁰Landscape Ecology Group, Institute of Biology and Environmental Sciences, University of Oldenburg, D-26111 Oldenburg, Germany. ¹¹Department of Systematic Botany and Functional Biodiversity, University of Leipzig, Johannisallee 21, 04103 Leipzig, Germany. ¹²AXA Chair in Biosphere and Climate Impacts, Grand Challenges in Ecosystems and the Environment and Grantham Institute – Climate Change and the Environment, Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot SL5 7PY, UK. ¹³Centre d'Ecologie Fonctionnelle et Evolutive (UMR 5175), CNRS-Université de Montpellier - Université Paul-Valéry Montpellier - EPHE, 34293 Montpellier Cedex 5, France. ¹⁴Plant Sciences (IBG-2), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany. ¹⁵Department of Forest Resources, University of Minnesota, St Paul, Minnesota 55108, USA. ¹⁶Hawkesbury Institute for the Environment, University of Western Sydney, Penrith New South Wales 2751, Australia. ¹⁷Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, UNSW Australia, Sydney, New South Wales 2052, Australia. ¹⁸Collections, The Royal Botanic Gardens Kew, Wakehurst Place, Ardingly, West Sussex, RH17 6TN, UK. ¹⁹Center for Biodiversity Management, P.O. Box 120, Yungaburra, Queensland 4884, Australia. ²⁰Department of Biological Sciences, George Washington University, Washington DC 20052, USA. ²¹Center for Conservation and Sustainable Development, Missouri Botanical Garden, St Louis, Missouri 63121, USA. ²²UMR 5174 Laboratoire Evolution et Diversité Biologique, CNRS & Université Paul Sabatier, Toulouse 31062, France. ²³Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancón, Panama. ²⁴Komarov Botanical Institute, Prof. Popov Street 2, St Petersburg 197376, Russia. ²⁵INRA, UMR1202 BIOGECO, F-33610 Cestas, France. ²⁶Université de Bordeaux, BIOGECO, UMR 1202, F-33600 Pessac, France. ²⁷International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Miami, Florida 33199, USA. ²⁸INRA, UMR Ecologie des Forêts de Guyane, 97310 Kourou, French Guiana. ²⁹Department of Theoretical and Applied Sciences, University of Insubria, Via J.H. Dunant 3, I-21100 Varese, Italy. ³⁰Department of Agricultural and Environmental Sciences (DiSAA), University of Milan, Via G. Celoria 2, I-20133 Milan, Italy. ³¹Département de biologie, Université de Sherbrooke, Sherbrooke, Québec J1K 2R1, Canada. ³²Biodiversity Informatics and Spatial Analysis, Jodrell Building, The Royal Botanic Gardens Kew, Richmond TW9 3AB, UK. ³³Unidad de Bioestadística, Centro Agronómico Tropical de Investigación y Enseñanza (CATIE), 7170 Turrialba, 30501, Costa Rica.

Table 1 | Range of variation in functional traits, geographic distribution and climatic conditions

	Abbreviation	Range	No. of species
Adult plant height (m)	H	0.001* to 90†	24,720
Stem specific density (mg mm ⁻³)	SSD	0.06‡ to 1.39§	11,356
Leaf area (mm ²)	LA	0.79* to 2.79 × 10 ⁶	12,173
Leaf mass per area (g m ⁻²)	LMA	4.9¶ to 1,507#	10,490
N content per unit leaf mass (mg g ⁻¹)	N _{mass}	2.48★ to 68.98**	8,695
Diaspore mass (mg)	SM	5.15 × 10 ⁻⁶ †† to 2.05 × 10 ⁷ ‡‡	24,779
Diaspore mass (mg) excluding pteridophytes	SM	3.0 × 10 ⁻⁴ §§ to 2.05 × 10 ⁷ ‡‡	24,449
Latitude (degrees)		55 S to 83.17 N	
Altitude (m)		-59 to 5,249	
Mean annual temperature (°C)		-27.22 to 29.97	
Mean annual sum of precipitation (mm yr ⁻¹)		<5 to 7,693	

Latitude and altitude are based on species occurrences in the Global Biodiversity Information Facility database (<http://www.gbif.org>). Mean annual temperature and annual sum of precipitation refer to CRU0.5 degree climatology. **Wolffia arrhiza* and *Azolla microphylla*; †*Sequoia sempervirens* and *Eucalyptus regnans*; ‡*Utricularia vulgaris*; §*Caesalpinia sclerocarpa*; ||*Victoria amazonica*; ¶*Myriophyllum aquaticum*; #*Agave americana*; ★*Hakea erecta*; ***Dipcadi glaucum*; ††*Blechnaceae*; ‡‡*Lodoicea maldivica*; §§*Laelia undulata* and *Alectra vogelii*.

how ecological filtering and evolution might further shape the Earth's plant trait space.

The trait space occupied by plants worldwide

Certain traits can be thought of as indexing positions of species along key dimensions of plant ecological strategy directly relevant to growth, survival and reproduction^{19,20,22,31–33}. We chose six traits whose fundamental importance for ecological strategy has been established unequivocally and for which data have recently become available for an unprecedented number of species worldwide³⁰. Among the six key traits (see Methods for details and references) adult plant height (H) corresponds with the ability to pre-empt light resources and disperse diaspores. Stem specific density (SSD) reflects a trade-off between growth potential and mortality risk from biomechanical or hydraulic failure. Leaf area (LA, size of an individual leaf) has important consequences for leaf energy and water balance. Leaf mass per area (LMA) and leaf nitrogen content per unit mass (N_{mass}) express different aspects of leaf strategy for resource capture and conservation: LMA reflects a trade-off between carbon gain and longevity, while N_{mass} reflects a trade-off between the benefits of photosynthetic potential and the costs of acquiring nitrogen and suffering herbivory. Diaspore mass (the mass of an individual dispersed seed or spore; SM) reflects a trade-off between seedling survival versus colonization ability in space and time. Ranges of trait variation span from 2 (SSD, N_{mass}) to 13 orders of magnitude (SM) (Table 1).

We investigated which portion of the six-dimensional trait space is occupied by vascular plants that now live on Earth. There are two primary reasons why plants might occupy a subset of the potential trait space: (1) values of independent traits are distributed along each axis in a clumped, non-uniform manner; and (2) there are inherent correlations between the values of different traits. We therefore built four null models varying the trait distributions and their correlation structure. We computed the volume of the six-dimensional convex hull³⁴, i.e. the smallest convex volume in hyperspace that contains the (log₁₀- and z-transformed) observed values of H, SSD, LA, LMA, N_{mass} and SM (for a visualization see <https://sdray.shinyapps.io/globalspectr/>; Supplementary Application 1), and compared it against hypervolumes from four null hypotheses (hv_{nm1} to hv_{nm4}; shown diagrammatically in Fig. 1 and described in detail in Methods). Hypervolumes hv_{nm1} to hv_{nm3} assume that the traits vary independently, resulting in a functional space spanning along six orthogonal axes. Null model 1 assumes that any combination of trait values can arise and escape natural selection with equal probability (for example ref. 35), thus extreme and central values are equally likely, each trait having a uniform distribution, and hv_{nm1} approximating a hypercube. Null model 2 assumes that extreme trait values are selected against during evolution and each trait has a log-normal distribution, with hv_{nm2}

approximating a hypersphere. Null model 3 imposes no assumptions about trait distributions but instead allows each trait to be distributed as observed and assumes traits are independent of one other. Null model 4 assumes that extreme values are selected against (i.e., log-normally distributed) and maintains the observed correlation structure among traits. Relative to null models 1 to 3, null model 4 collapses the multidimensional trait-space occupied by plants (hv_{nm4}) into an elongated hyperellipsoid.

We found that the observed hypervolume (hv_{obs}) is much smaller than hypervolumes expected under the first three null models (hv_{nm1} to hv_{nm3}) (Fig. 1). While closer in size to hv_{nm4}, it is still 20% smaller. It also shows greater aggregation of species ('lumpiness') in multivariate

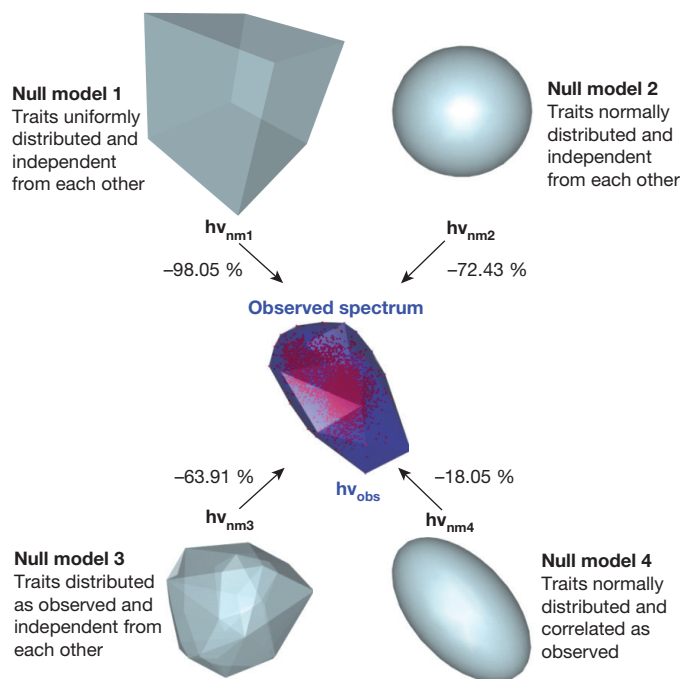


Figure 1 | The volume in trait space occupied by vascular plant species is strongly constrained compared to theoretical null models. The five diagrams are pictorial representations based on three out of the six trait dimensions forming the hypervolumes under scrutiny. The hypervolumes are constructed on the basis of log₁₀- and z-transformed observed values of H, SSD, LA, LMA, N_{mass} and SM (observed hypervolume = hv_{obs}), or on the bases of four different null models of multivariate variation of those traits (hv_{nm1} to hv_{nm4}) (see Methods). Numbers adjacent to arrows indicate percentage reductions in size of hv_{obs} compared to the null-model hypervolumes (all significant at $P < 0.001$).

space than expected under each of the null models (Supplementary Table 1). Thus the restriction of the observed hypervolume mainly reflects correlations among the six traits, and also—to a smaller degree—a greater concentration than expected under multivariate normality. In sum, the trait hypervolume occupied by living vascular plants is strongly constrained, converging towards a relatively small set of successful trait combinations.

Main trends of variation

Within the observed worldwide plant trait space we identified the main independent dimensions of variation. Seventy-four percent of the variation in the six-dimensional space was accounted for by the plane defined by the first two principal components (PC), the only PC found to contain significant, non-redundant information (Fig. 2, Extended Data Table 1 and Extended Data Fig. 2; all PC displayed at <https://sdray.shinyapps.io/globalspectrPC/>; Supplementary Application 2). Within this plane two notable dimensions of trait variation stand out. One dimension runs from short species tending to have small diaspores to tall species tending to have large diaspores (lower left to upper right in Fig. 2a, 'H-SM', more strongly associated with PC1 than PC2). The other (upper left to lower right in Fig. 2a, 'LMA- N_{mass} ', more strongly associated with PC2 than PC1) runs from species with

cheaply constructed, 'acquisitive' leaves (low-LMA, nitrogen-rich) to species with 'conservative' leaves (high-LMA, nitrogen-poor) that are expected to have longer leaf lifespan and higher survival in the face of abiotic and biotic hazards^{7,10,36}. Stem specific density (SSD) and leaf area (LA) also load heavily on the plane and are correlated with both the H-SM and the LMA- N_{mass} dimensions (Fig. 2a, Extended Data Table 1). Although SSD and SM increase with plant height, at any given H there is considerable independent variation in both (Extended Data Fig. 3a, f), and at any given LMA and N_{mass} there is considerable independent variation in LA (Extended Data Figs 3b, c). These general patterns are robust (Extended Data Table 1) with respect to species selection (for example, considering angiosperms rather than all species), to re-running analyses on a 45,507-species 'gap-filled' trait matrix rather than the 2,214-species six-trait matrix, and to exclusion of individual traits (for example, using only one rather than both leaf economic traits). The outer reaches of the main plane of variation represent extreme combinations of plant size and leaf structure and function (see circled numbers in Fig. 2a, and Extended Data Table 2 for illustrative species), with a wide gradient of intermediate trait combinations between them, together expressing the rich variety of ways in which plants balance the challenges of growth, survival and reproduction.

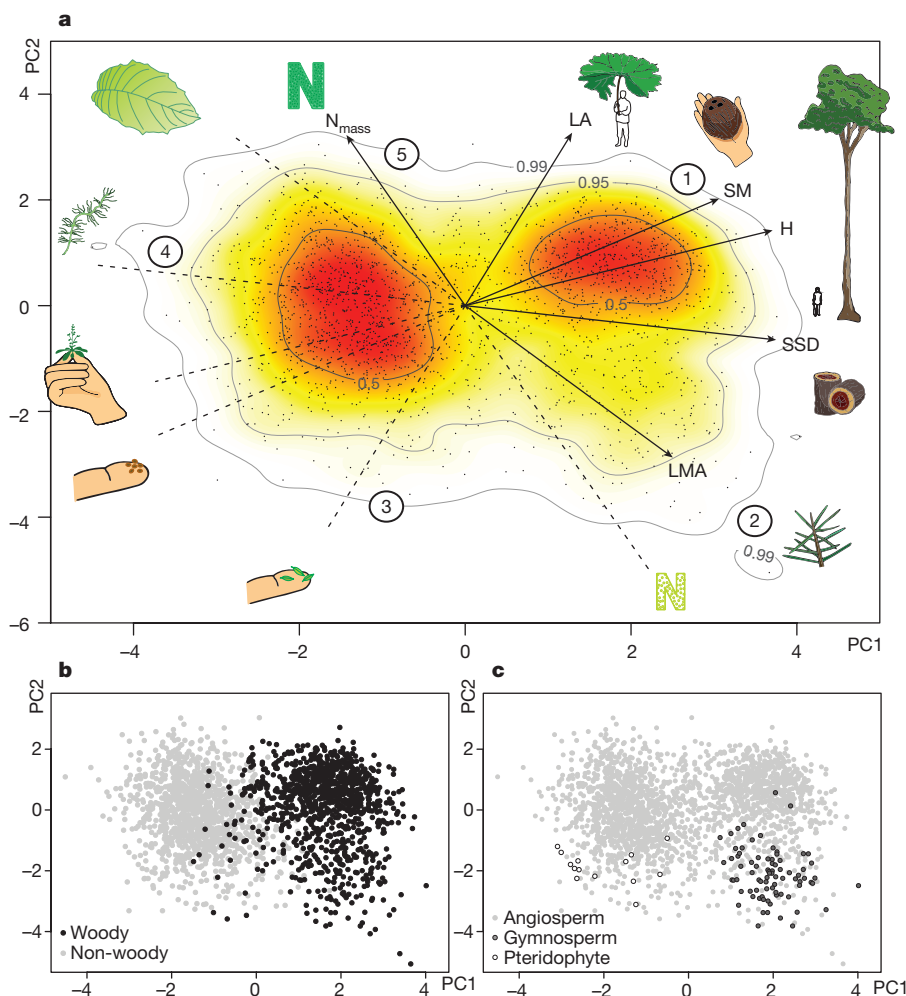


Figure 2 | The global spectrum of plant form and function. **a**, Projection of global vascular plant species (dots) on the plane defined by principal component axes (PC) 1 and 2 (details in Extended Data Table 1 and Extended Data Fig. 2). Solid arrows indicate direction and weighing of vectors representing the six traits considered; icons illustrate low and high extremes of each trait vector. Circled numbers indicate approximate position of extreme poles of whole-plant specialization, illustrated by typical species (Extended Data Table 2). The colour gradient indicates

regions of highest (red) to lowest (white) occurrence probability of species in the trait space defined by PC1 and PC2, with contour lines indicating 0.5, 0.95 and 0.99 quantiles (see Methods, kernel density estimation). Red regions falling within the limits of the 0.50 occurrence probability correspond to the functional hotspots referred to in main text. **b**, **c**, location of different growth-forms (**b**) and major taxa (**c**) in the global spectrum.

Major taxa, growth-forms, and functional hotspots

Different plant groups distribute unevenly in the global spectrum of form and function. Both herbaceous and woody growth-forms show considerable variation along the two main dimensions (Fig. 2b). The two groups are offset along the H–SM dimension (Fig. 2b), with woody species, on average, being taller and having larger seeds and leaves; woody species also tend to have higher SSD and LMA than herbaceous species (Extended Data Fig. 3a–e). Also, although taller species have larger seeds in both herbaceous and woody species-groups, the relationship is only very weak in herbaceous species (Extended Data Figs 3f and 4). In sum, the distinction in traits between herbaceous and woody growth-forms goes beyond the obvious difference in height and stem structure that has been recognized since antiquity³⁷. At the same time, there exist commonalities in trait coordination and trade-offs across both herbaceous and woody plants, shown here at a global scale for the first time. For example, herbaceous and woody plants overlap widely along the LMA– N_{mass} dimension (Fig. 2b), particularly in regard to N_{mass} (Extended Data Fig. 3c), and LMA and N_{mass} are largely independent from LA in both groups (Extended Data Fig. 3b, c). Further, while neither SSD nor LMA increases with plant stature within either group (Extended Data Figs 3a, e and 4), LA increases with H in both (Extended Data Fig. 3d). These multivariate trends are summarised by the clear distinction of herbaceous and woody species-groups along PC1, and their broad overlap along PC2 (Extended Data Fig. 2a).

There are also strong differences in trait-space occupancy by major taxa. For gymnosperms, high costs of seed packaging and abortion are thought to set a lower bound on seed size^{38,39}. Accordingly in Fig. 2c gymnosperms are confined to the right hand side (see also Extended Data Fig. 2b, and, for examples, Extended Data Table 2). The emergence of angiosperms allowed a considerable extension into smaller seed size⁴⁰ that is manifest in extant species. This also opened up lifestyles involving colonization of open ground, shorter lifespans and herbaceous growth-form (towards the left of Fig. 2a). The second major angiosperm innovation whose footprint is evident in the trait space concerns xylem vessels. Angiosperm vessels are longer and larger-diameter conduits than gymnosperm and pteridophyte tracheids, permitting much higher hydraulic conductivities. This, together with a greater density of leaf veins, has allowed angiosperms to deliver a faster transpiration stream while requiring less volume within the leaf⁴¹. These anatomical innovations have made it possible for angiosperms to extend the range of leaf stomatal conductances and photosynthetic capacities to higher values (requiring coordinated higher N_{mass}) and the range of LMA to lower values compared to gymnosperms and pteridophytes (Fig. 2c). Higher hydraulic conductivity presumably also enabled the evolution of very large leaves in angiosperms, and a far wider variety in leaf morphology too. Nevertheless, while angiosperm innovations have expanded trait space considerably towards higher leaf N_{mass} and LA and (compared with gymnosperms) lower diaspore mass, angiosperms also converged on the same zones of trait space as gymnosperms and pteridophytes, as seen in the lower right and lower left of the global trait plane (Fig. 2c).

There are two clear functional hotspots—areas of particularly dense species occupation—in trait space (Fig. 2a). The bimodality resides in H and in SSD, rather than in LMA, N_{mass} or LA (Extended Data Fig. 4). The first hotspot almost entirely corresponds to herbaceous plants and lies at the core of the distribution of both graminoid (grass-like) and non-graminoid herbs, having small, acquisitive leaves and small seeds. The second hotspot lies within the trait space occupied by woody plants, positioned towards the upper right corner of Fig. 2a. It consists mostly of tree species of moderate to great height with large leaves and large seeds; plants other than angiosperms are almost completely absent from it. Many phylogenetically distant families and orders of angiosperms are represented within each hotspot (Supplementary Table 2), indicating that these prevalent ecological trait constellations represent successful solutions acquired repeatedly through the evolutionary history of vascular plants.

Discussion

Our findings show that the trait space currently occupied by vascular plants is quite restricted compared to the range of possibilities that would exist if traits varied independently. Importantly, this finding arises from the combined analysis of six traits describing different plant organs, and from a wider spread of taxa and life histories than has previously been possible. It yields the most comprehensive picture to date of how the remarkable functional diversity of vascular plants seen on Earth today has been able to evolve within very general constraints. This worldwide functional six-trait space is wide, diverse and lumpy, with its fringes occupied by species (indicated with circled numbers in Fig. 2a) ranging from the short model plant thale cress (*Arabidopsis thaliana*) to the 60-m tall Brazil Nut tree (*Bertholletia excelsa*), from flimsy watermilfoil (*Myriophyllum spicatum*) to tough monkey puzzle tree (*Araucaria araucana*), from the tender but toxic devil's snare (*Datura stramonium*) to the hardy needlewood (*Hakea leucoptera*), from the minute leaves and seeds of heather (*Calluna vulgaris*) to the large leaves and seeds of lotus (*Nelumbo nucifera*) (description and additional illustrative species in Extended Data Table 2). Yet, this variation of the six key traits is largely concentrated into a plane.

Stem density, leaf size and diaspore size represent trade-offs within distinct aspects of plant biology³², and in previous studies of trends across different plant organs, these traits have shown considerable variation that is independent from whole-plant size and leaf carbon economy. However, those analyses were based on far more restricted datasets than considered here, in terms of growth-forms, habitats, or both, considering for example tropical woody species^{21–23,26}, temperate semiarid pine forests²⁴, or countrywide herbaceous floras²⁷. At the global scale of our study, these three traits do not constitute major independent dimensions; rather, substantial variation in them is captured by the plane that summarizes global variation in vascular plant design (Fig. 2). Our results are correlative and cannot prove rigorously why such a large share of the potential trait volume is not occupied. Still, from first principles many more combinations of traits than those observed seem feasible as far as biomechanics and evolutionary genetics are concerned. We suggest the concentration into two dimensions and the lumpiness within that plane reflect the major trait constellations that are competent and competitive across the ecological situations available on Earth today.

The patterns described here pertain to fundamental aspects of form and function critical to growth, survival and reproduction of the vast majority of vascular plants on Earth. Importantly, plants converge and diverge in many more ways than explored here, through variation in a vast array of traits beyond the scope of our analysis, related to the fine-tuning of different taxa to specific abiotic and biotic conditions in their habitat (for example refs 42–44). Such variation fits within the more general patterns shown here.

More broadly, our findings are directly relevant to a number of long-running and emerging broad-scale scientific initiatives. First, our findings provide the widest empirical context so far for examining theories that have focused on plant ecological strategies—on different aspects of the Darwinian struggle for existence. For example, the H–SM dimension could be seen as reflecting the r (colonization) versus K (exploitation) continuum^{4,45}. The LMA– N_{mass} dimension reflects the A (adversity-selection) continuum^{4,46}, acquisitive-conservative continuum^{5,7,19,20} or leaf economic spectrum¹⁰. The positions signalled by numbers 3, 4, and 5 in Fig. 2a (and described in Extended Data Table 2) could arguably roughly correspond to the stress-tolerant, ruderal, and competitor strategies of Grime^{3,11,19}. Interestingly, the functional hotspots lie at intermediate positions on the plane rather than at any of these extreme positions (that is, r versus K, acquisitive versus conservative resource economy, C, S or R-strategy).

Second, the global spectrum we describe has potential to improve emerging large-scale vegetation and ecosystem models (for example see refs 47–49) because we clearly show (Fig. 2a and Extended

Data Fig. 4) that only a limited set of combinations are observed from six plant traits most fundamental to survival, growth and reproduction.

More generally, our findings—as encapsulated in the plane of Fig. 2—establish a backdrop against which many other facets of plant biology can be placed into a broader context. Plant lineages, evolutionary trajectories, and historical and contemporary plant communities and biomes can be mapped onto this global trait spectrum. Trait variation in any given physical setting can be compared to the worldwide background. Model species such as *Arabidopsis thaliana* (located at one extreme of the spectrum) can be positioned against this backdrop, helping to judge how typical or otherwise their physiology and natural history might be. The global spectrum of plant form and function is thus, in a sense, a galactic plane within which we can position any plant—from star anise to sunflower—based on its traits.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 September; accepted 26 November 2015.

Published online 23 December 2015.

- Schimper, A. *Plant-geography upon a Physiological Basis* (Oxford Univ. Press, 1903).
- Warming, E. *Ecology of Plants* (Clarendon, 1909).
- Grime, J. P. Vegetation classification by reference to strategies. *Nature* **250**, 26–31 (1974).
- Southwood, T. R. E. Habitat, the templet for ecological strategies? *J. Anim. Ecol.* **46**, 337–365 (1977).
- Chapin, F. S. The mineral nutrition of wild plants. *Annu. Rev. Ecol. Syst.* **11**, 233–260 (1980).
- Niklas, K. J. *The Evolutionary Biology of Plants* (Univ. of Chicago Press, 1997).
- Reich, P. B., Walters, M. B. & Ellsworth, D. S. From tropics to tundra: global convergence in plant functioning. *Proc. Natl Acad. Sci. USA* **94**, 13730–13734 (1997).
- Ackerly, D. Functional strategies of chaparral shrubs in relation to seasonal water deficit and disturbance. *Ecol. Monogr.* **74**, 25–44 (2004).
- Field, C. H. & Mooney, H. A. In *On the Economy of Plant Form and Function* (ed. Givnish, T. J.) pp. 25–49 (Cambridge Univ. Press, 1986).
- Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).
- Pierce, S., Brusa, G., Vagge, I. & Cerabolini, B. E. L. Allocating CSR plant functional types: the use of leaf economics and size traits to classify woody and herbaceous vascular plants. *Funct. Ecol.* **27**, 1002–1010 (2013).
- Li, L. *et al.* Leaf economics and hydraulic traits are decoupled in five species-rich tropical-subtropical forests. *Ecol. Lett.* **18**, 899–906 (2015).
- Chave, J. *et al.* Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12**, 351–366 (2009).
- Zanne, A. E. *et al.* Angiosperm wood structure: Global patterns in vessel anatomy and their relation to wood density and potential conductivity. *Am. J. Bot.* **97**, 207–215 (2010).
- Salisbury, E. Seed size and mass in relation to environment. *Proc. R. Soc. B* **186**, 83–88 (1974).
- Thompson, K. Seeds and Seed Banks. *New Phytol.* **106**, 23–34 (1987).
- Moles, A. T., Falster, D. S., Leishman, M. R. & Westoby, M. Small-seeded species produce more seeds per square metre of canopy per year, but not per individual per lifetime. *J. Ecol.* **92**, 384–396 (2004).
- Lambers, H. & Poorter, H. Inherent variation in growth rate between higher plants: a search for physiological causes and ecological consequences. *Adv. Ecol. Res.* **23**, 187–261 (1992).
- Grime, J. P. *et al.* Integrated screening validates primary axes of specialisation in plants. *Oikos* **79**, 259–281 (1997).
- Diaz, S. *et al.* The plant traits that drive ecosystems: evidence from three continents. *J. Veg. Sci.* **15**, 295–304 (2004).
- Wright, I. J. *et al.* Relationships among ecologically important dimensions of plant trait variation in seven neotropical forests. *Ann. Bot.* **99**, 1003–1015 (2007).
- Poorter, L. *et al.* Are functional traits good predictors of demographic rates? Evidence from five neotropical forests. *Ecology* **89**, 1908–1920 (2008).
- Baraloto, C. *et al.* Decoupled leaf and stem economics in rain forest trees. *Ecol. Lett.* **13**, 1338–1347 (2010).
- Laughlin, D. C., Leppert, J. J., Moore, M. M. & Sieg, C. H. A multi-trait test of the leaf-height-seed plant strategy scheme with 133 species from a pine forest flora. *Funct. Ecol.* **24**, 493–501 (2010).
- Freschet, G. T., Cornelissen, J. H. C., Van Logtestijn, R. S. P. & Aerts, R. Evidence of the ‘plant economics spectrum’ in a subarctic flora. *J. Ecol.* **98**, 362–373 (2010).
- Patiño, S. *et al.* Coordination of physiological and structural traits in Amazon forest trees. *Biogeosciences* **9**, 775–801 (2012).
- Pierce, S., Bottinelli, A., Bassani, I., Ceriani, R. M. & Cerabolini, B. E. L. How well do seed production traits correlate with leaf traits, whole-plant traits and plant ecological strategies? *Plant Ecol.* **215**, 1351–1359 (2014).
- Price, C. A. *et al.* Are leaf functional traits ‘invariant’ with plant size and what is ‘invariance’ anyway? *Funct. Ecol.* **28**, 1330–1343 (2014).
- Cornwell, W. K. *et al.* Functional distinctiveness of major plant lineages. *J. Ecol.* **102**, 345–356 (2014).
- Kattge, J. *et al.* TRY—a global database of plant traits. *Glob. Change Biol.* **17**, 2905–2935 (2011).
- Weiher, E. *et al.* Challenging Theophrastus: a common core list of plant traits for functional ecology. *J. Veg. Sci.* **10**, 609–620 (1999).
- Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. A. & Wright, I. J. Plant ecological strategies: some leading dimensions of variation between species. *Annu. Rev. Ecol. Syst.* **33**, 125–159 (2002).
- Garnier, E., Navas, M. & Grigulis, K. *Plant Functional Diversity - Organism Traits, Community Structure, and Ecosystem Properties* (Oxford Univ. Press, 2016).
- Cornwell, W. K., Schwilik, L. D. & Ackerly, D. D. A trait-based test for habitat filtering: convex hull volume. *Ecology* **87**, 1465–1471 (2006).
- Leimar, O. Evolutionary change and Darwinian demons. *Selection* **2**, 65–72 (2001).
- Coley, P. D., Bryant, J. P. & Chapin, F. S. III. Resource availability and plant antiherbivore defense. *Science* **230**, 895–899 (1985).
- Theophrastus. *Enquiry into Plants and Minor Works on Odours and Weather Signs* (translated by Hort, A. F.) (Heinemann, 1916).
- Haig, D. & Westoby, M. Seed size, pollination costs and angiosperm success. *Evol. Ecol.* **5**, 231–247 (1991).
- Lord, J. M. & Westoby, M. Accessory costs of seed production and the evolution of angiosperms. *Evolution* **66**, 200–210 (2012).
- Eriksson, O., Friis, E. M. & Löfgren, P. Seed size, fruit size and dispersal spectra in angiosperms from the Early Cretaceous to the Late Tertiary. *Am. Nat.* **156**, 47–58 (2000).
- Boyce, C. K., Brodribb, T. J., Feild, T. S. & Zwieniecki, M. A. Angiosperm leaf vein evolution was physiologically and environmentally transformative. *Proc. R. Soc. B* **276**, 1771–1776 (2009).
- Grubb, P. J. The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biol. Rev. Camb. Philos. Soc.* **52**, 107–145 (1977).
- Grubb, P. J. In *The Population Structure of Vegetation* (ed. White, J.) pp. 595–621 (Springer, 1985).
- Grime, J. P. Declining plant diversity: empty niches or functional shifts? *J. Veg. Sci.* **13**, 457–460 (2002).
- MacArthur, R. H. & Wilson, E. O. *The Theory of Island Biogeography* (Princeton Univ. Press, 1967).
- Greenslade, P. Adversity selection and the habitat templet. *Am. Nat.* **122**, 352–365 (1983).
- van Bodegom, P. M., Douma, J. C. & Verheijen, L. M. A fully traits-based approach to modeling global vegetation distribution. *Proc. Natl Acad. Sci. USA* **111**, 13733–13738 (2014).
- Scheiter, S., Langan, L. & Higgins, S. I. Next-generation dynamic global vegetation models: learning from community ecology. *New Phytol.* **198**, 957–969 (2013).
- Harfoot, M. B. J. *et al.* Emergent global patterns of ecosystem structure and function from a mechanistic general ecosystem model. *PLoS Biol.* **12**, e1001841 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to the many researchers who contributed to this study by making their data available, helping to check information, and/or providing comments at various stages. The study was supported by the TRY initiative on plant traits (<http://www.try-db.org>). The TRY database is hosted at the Max Planck Institute for Biogeochemistry (Jena, Germany) and supported by DIVERSITAS/Future Earth, the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, and BACI (grant ID 640176). The study has also been partially supported by the following institutions and grants to S.D.: Universidad Nacional de Córdoba and CONICET, FONCYT (PICT 554) and SECyT (Argentina), The Leverhulme Trust, UK, and Inter-American Institute for Global Change Research (IAI) SGP-CRA2015 (supported by US National Science Foundation grant GEO-1138881).

Author Contributions Order in list of authors reflects overall participation in this article. S.D., J.K. and S.L. designed the study. S.D., J.K., J.H.C.C., I.J.W., S.L., M.K., C.W., E.G., I.C.P., M.W., H.P., P.B.R., A.T.M., J.D., A.N.G., A.E.Z., J.C., S.J.W., S.N.S., H.J., C.B., B.C., S.P., B. S. and D.K. contributed substantial amounts of data. S.D., J.K., G.B., A.G. and V.F. constructed the database. S.D., J.K., J.H.C.C., I.J.W., S.L., S.Dr., B.R., M.K., C.W., E.G., F.C., J.S.J., N.R., M.D.M. and L.D.G. carried out analyses. S.D., J.K., J.H.C.C., I.J.W., S.L., M.K., C.W., I.C.P., M.W. and P.B.R. wrote the article with contributions from S.Dr., B.R., E.G., H.P., A.T.M., J.D., A.N.G., A.E.Z., J.C., S.J.W., S.N.S., H.J., C.B., B.C., S.P., B.S., D.K., F.C., M.D.M. and L.D.G.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.D. (sandra.diaz@unc.edu.ar).

METHODS

Plant trait definitions and ecological meaning. Adult plant height (typical height of the upper boundary of the main photosynthetic tissues at maturity; hereafter *H*) is the most common measure of whole plant size and indicates ability to pre-empt resources, including the ability of taller plants to display their leaves above those of others and therefore outcompete them; it also relates to whole plant fecundity and facilitation of seed dispersal^{32,50–53}. Taller plants intercept more light but, trading off against that, construction and maintenance costs and risk of breakage increase with height^{22,54}. Large stature has been repeatedly selected for in different lineages during the evolution of land plants, although achieved very differently in different clades⁵⁵.

Stem specific density (dry mass per unit of fresh stem volume; SSD) is a second key index of construction costs and structural strength. Although SSD is more commonly measured on trees, here we used data for both herbaceous and woody species. At least among woody species, stem specific density is positively linked with plant mechanical strength, hydraulic safety and resistance to biotic agents^{13,14,56}. In high-precipitation systems wood density underpins a successional continuum running from low-SSD, fast-growing, light-demanding species to high-SSD, slow-growing, shade-tolerant species. More broadly SSD characterizes a trade-off between fast growth with high mortality rates versus slow growth with high survival^{25,57}.

Leaf area (one-sided surface area of an individual lamina; *LA*) is the most common measure of leaf size. It is relevant for light interception and has important consequences for leaf energy and water balance^{58–60}. *LA* affects leaf temperature via boundary layer effects. The larger the lamina, the lower the heat exchange, the diffusion of carbon dioxide and water vapour per unit leaf area between a leaf and the surrounding air. *LA* is known to be constrained by climatic and microclimatic conditions and also by the allometric consequences of plant size, anatomy and architecture^{61–63}.

Leaf dry mass per unit of lamina surface area (leaf mass per area; *LMA*) and leaf nitrogen content per unit of lamina dry mass (N_{mass}) reflect different aspects of leaf-level carbon-gain strategies, in particular the ‘leaf economic spectrum’ running from ‘conservative’ species with physically robust, high-*LMA* leaves with high construction costs per unit leaf area and long expected leaf lifespan (and thus duration of photosynthetic income) to ‘acquisitive’ species with tender, low-cost low-*LMA* leaves, and short leaf lifespan^{5,9,10,18–20,32,64,65}. *LMA* relates the area of light interception to leaf biomass, being an expression of how much carbon is invested per unit of light-intercepting area, and thus reflecting a trade-off between construction cost and longevity of lamina. N_{mass} is directly related to photosynthesis and respiration, as the majority of leaf nitrogen is found in metabolically active proteins. N_{mass} reflects a trade-off between, on the one hand, two different costs that increase with higher N_{mass} (to acquire N, and potentially suffer more herbivory) and, on the other hand, the greater photosynthetic potential that higher N_{mass} allows.

Diaspore mass (mass of an individual seed or spore plus any additional structures that assist dispersal and do not easily detach; *SM*) indexes species along a dimension describing the trade-off between seedling competitiveness and survival on the one hand, and dispersal and colonization ability on the other^{16,66–68}. As a broad generalization small seeds can be produced in larger numbers with the same reproductive effort and, at a given plant height, be dispersed further away from the parent plant and form persistent seed banks, whereas large seeds facilitate survival through the early stages of recruitment, and higher establishment in the face of environmental hazards (for example deep shade, drought, herbivory)^{32,69–72}.

Dataset description. We compiled a global dataset containing 46,085 species and 601,973 cells, of which 92,212 correspond to quantitative species-level trait information, based on > 800,000 trait measurements for the six traits of interest on > 500,000 plant individuals. The vast majority of data were compiled from pre-existing smaller datasets contributed to the TRY Plant Trait Database³⁰ (<https://www.try-db.org>, accessed May 2015). The dataset was supplemented by published data not included in TRY and a small number of original unpublished data by W. J. Bond, J.H.C.C., S.Di., L. Enrico, M. T. Fernandez-Piedade, L.D.G., D.K., M.K., N. Salinas, E.-D. Schulze, K. Thompson, and R. Urrutia. The final dataset (BLOB) was derived from 175 studies^{10,11,13,20,21,23–25,56,57,64,65,73–235}.

In this global analysis, each species, identified subspecies or local variety is represented by a single value for each trait. This value is the geometric mean of all the observations of a trait present in the TRY Plant Trait Database and additional databases incorporated to the present dataset. The number of observations per trait and species range from a single one (in the case of rare, geographically restricted species) to hundreds (in the case of common species with wide geographical distribution). In this way, the analysis incorporates the high intraspecific variation that is sometimes observed in widespread species^{30,185,236–238}. In addition and fully acknowledging their existence, intraspecific variations are assumed to

be comparatively small in the context of the vast range of variation contained in this worldwide dataset³⁰.

Species names were standardized and attributed to families according to The Plant List (<http://www.theplantlist.org/>; accessed 2015). Attribution of families to higher-rank groups was made according to APG III (2009) (<http://www.mobot.org/MOBOT/research/APweb/>). Information about primary growth-form (woodiness, woody, semi-woody, non-woody) and secondary growth-form (herbaceous non-graminoid, herbaceous graminoid, herbaceous non-graminoid/shrub, shrub, shrub/tree, tree, climber, succulent, other) was added based on a look-up table of categorical plant-traits³⁰ (<https://www.try-db.org/TryWeb/Data.php#3>) and additional information from various sources; >86% species were allocated to categories according to primary growth-form, and >80% according to secondary growth-form.

Species distribution data were derived from the Global Biodiversity Information System (GBIF; <http://www.gbif.org>) and combined with 0.5 × 0.5 degree gridded long term climate information derived from CRU (<http://www.cru.uea.ac.uk/data>). **Trait measurement.** In the case of published datasets, trait measurement methods are in the original publications listed in Dataset description. In the case of unpublished records, they were measured following the protocols specified in the context of the LEDA project (<https://www.leda-traitbase.org>) or in ref. 239 (<http://www.nucleodiversus.org/index.php?mod=page&id=79>). All data were unit-standardized and subjected to error detection and quality control (see below).

Treatment of pteridophyte spore mass. The trait values for diaspore mass of pteridophytes were estimated based on literature data for spore radius (*r*). We made crude assumptions that spores would be broadly spherical, with volume = $(4/3)\pi r^3$, and that their density would be 0.5 mg mm⁻³. Although these assumptions were clearly imprecise, we are confident they result in spore masses within the right order of magnitude (and several orders of magnitude smaller than seed mass of spermatophytes). Most data were from ref. 240, data for *Sadleria pallida* were from ref. 241, for *Pteridium aquilinum* from ref. 242, and for *Diphysastrum* spp. from ref. 243.

Treatment of stem specific density in herbaceous species. Data on stem specific density (SSD) are available for a very large number of woody species, but only for very few herbaceous species. To incorporate this fundamental trait in our analysis, we complemented SSD of herbaceous species using an estimation based on leaf dry matter content (LDMC), a much more widely available trait, and its close correlation to stem dry matter content (StDMC, the ratio of stem dry mass to stem water-saturated fresh mass). StDMC is a good proxy of SSD in herbaceous plants with a ratio of approximately 1:1 (ref. 100), despite substantial differences in stem anatomy among botanical families²⁴⁴, including those between non-monocotyledons and monocotyledons (where sheaths were measured). We used a data set of 422 herbaceous species collected in the field across Europe and Israel, and belonging to 31 botanical families¹⁴⁴ to parameterize linear relationships of StDMC to LDMC. The slopes of the relationship were significantly higher for monocotyledons than for other angiosperms ($F = 12.3$; $P < 0.001$); within non-monocotyledons, the slope for Leguminosae was higher than that for species from other families. We thus used three different equations to predict SSD for 1963 herbaceous species for which LDMC values were available in TRY: one for monocotyledons ($\text{SSD} = 0.888 \times \text{LDMC} + 2.69$), one for Leguminosae ($\text{SSD} = 0.692 \times \text{LDMC} + 47.65$), and a third one for other non-monocotyledons ($\text{SSD} = 0.524 \times \text{LDMC} + 95.87$).

Error detection and data quality control. The curation of the dataset faced a double challenge: (1) detecting erroneous entries (due to errors in sampling, measurement, unit conversion, etc.); and (2) ensuring that extreme values that correspond to truly extreme values of traits in nature are not mistakenly identified as outliers and therefore excluded from the dataset. To deal with these challenges, we took the following approach: Trait records measured on juvenile plants and on plants grown under non-natural environmental conditions were excluded from the dataset. Duplicate trait records (same species, similar trait values, no information on different measurement locations or dates) and obvious errors (for example $\text{LMA} < 0$) were excluded from the dataset. We then identified potential outliers following the approach described in ref. 30. Trait records with a distance of >4 standard deviations from the mean of species, genus, family or higher-rank taxonomic group were excluded from the dataset unless their retention could be justified from external sources. Trait records with a distance of >3 standard deviations from the mean of species, genus, family or phylogenetic group were identified, checked by domain experts for plausibility and retained or excluded accordingly. The remaining dataset was used to calculate species mean trait values. Finally, the ten most extreme species mean values of each trait (smallest and largest) were again checked for reliability against external sources.

Construction of observed and simulated six-trait convex hull hypervolumes. In order to explore the constraints underlying the trait space occupied by species in our dataset, we used the convex hull approach of ref. 34, which has been applied

successfully to a wide range of datasets, including disjoint ones²⁴⁵. The application of a recently developed—and therefore less widely tested—method proposed for “holey” datasets²⁴⁵ yielded similar results.

We computed a six-dimensional convex hull volume (i.e. a six-dimensional measure of the minimum convex volume of trait space occupied by species in our dataset, hereafter Hv_{obs}) on the basis of the observed values of H , SSD , LA , LMA , N_{mass} and SM , and compared it to four null model volumes (h_{vnm1-4}) constructed under four different sets of assumptions. In all four cases the null hypothesis was $H_0: Hv_{obs} = h_{vnm}$ and the alternative hypothesis was $H_1: Hv_{obs} < h_{vnm}$ (‘the volume of the convex hull defined by the observed species is smaller than the volume occupied by species if their traits were generated under the null hypothesis’). Observed data were \log_{10} -transformed and standardized to zero mean and unit variance (z -transformation). Percentages in Fig. 1 indicate the reduction in size of the observed hypervolume compared to the average of 999 hypervolumes generated from the assumptions of each null model (Monte-Carlo permutations²⁴⁶). To control for outliers, computation of volumes were performed on the observed and simulated convex hulls containing 95% of species located closest to the centroid. A visualization of the observed dataset and the four null models in three-dimensional trait spaces is available at <https://sdray.shinyapps.io/globalspectr/> (Supplementary Application 1). The R script used for hypervolume computation is provided at ftp://pbil.univ-lyon1.fr/pub/datasets/drays/Diaz_Nature/.

Null model 1. Species traits vary independently and each of them comes from a uniform distribution. This null model assumes that each of the six traits represents an independent axis of specialization (i.e. the traits define a six-dimensional manifold) and that the occurrence of extreme and central values is equally probable. This uniform independent trait distribution represents a “Darwinian Demon”³⁵ scenario, where any combination of trait values can arise from mutation and escape from the natural selection process with equal probability. Simulated data are generated by randomly and independently sampling from independent uniform distributions whose range limits are constrained to the 0.025 and 0.975 quantiles of the observed trait values. The shape of the hypervolume under this null model (h_{vnm1}) is a hypercube.

Null model 2. Species traits vary independently and each of them comes from a normal distribution. This null model assumes that all six traits evolve independently, as in null model 1. However, extreme trait values are selected against during evolution. Simulated data were obtained by randomly and independently selecting from univariate normal distributions with standard deviation determined by the transformed observed data. The corresponding hypervolume (h_{vnm2}) is a hypersphere.

Null model 3. Species traits vary independently but—unlike in the previous models—there is no assumption about the distribution of trait variation; each trait varies according to the observed univariate distributions. Simulated data were obtained by permuting the values for each trait independently and therefore destroying the covariance amongst traits. Under this null hypothesis (h_{vnm3}) the hypervolume can take many potential shapes, emerging from the possible combinations of independently sampled plant trait observations.

Null model 4. Species traits are normally distributed and follow the estimated correlation structure of the observed dataset. This null model assumes that there are less than six independent axes of specialization and that extreme values are selected against. Simulated data were obtained by generating multivariate normal variables with standard deviations of the transformed observed data using the correlation structure of the observed dataset. The corresponding hypervolume (h_{vnm4}) is a hyperellipsoid. Deviations of observed data from null model 4 can be explained by deviations of the transformed observed univariate distributions from normal distributions, either showing lower tails than those expected in a normal distribution or by the non-observation of some combinations of extreme trait values, leading to truncated distributions, or by bimodal distributions.

Test for concentration of species within the observed convex hull. For each trait, values were partitioned in 10 bins so that the multivariate space was divided in 10^6 cells. The number of species per cell was computed and cumulative frequency curves were built for observed data and null models. For each null model, we simulated 999 datasets and computed the 0.025–0.975 interquartile range and the median. We then determined N_{10} and N_{50} , the minimum number of cells needed to cover 10% or 50% of species.

Principal component analysis (PCA). We performed PCAs on different versions of the observed dataset and a gap-filled version using the statistical software package InfoStat²⁴⁷ and the R-function ‘princomp’. Again all analyses were carried out on the correlation matrix of \log_{10} -transformed variables (traits), which is equivalent to using standardized data (z -transformation), which is considered appropriate for data with different measurement scales²⁴⁸. The number of significant PCA axes to be retained in order to minimize both redundancy and loss of information was determined using the procedure proposed by ref. 249, which allows one to test the significance of dimensionality in PCA. A visualization of the space occupied by

vascular plants in the space defined by all six PCA axes (three at a time) is available at <https://sdray.shinyapps.io/globalspectrPC/> (Supplementary Application 2).

Differences in the position of different major taxa and growth forms along PC1 and PC2 were tested using analysis of variance (Extended Data Fig. 2). Because of the large number of data, we used an alpha level of 0.01 to reject the null hypothesis. ANOVA was carried out using a linear mixed model to take into account the lack of homoscedasticity due to different group sizes. We used AIC and BIC criteria to select the best model considering heterogeneous variances. When the ANOVA null hypothesis was rejected, means were compared using Fisher's least significant difference ($P = 0.01$). Data were analysed using the lme function of the nlme²⁵⁰ and lme4²⁵¹ R-packages²⁵² interfaced by InfoStat Statistical Software version 2015 (ref. 247).

Test for robustness and representativeness of multivariate analysis results. In order to test if results shown in Fig. 2 and Extended Data Table 1 were robust with respect to the selection of traits and species and representative for vascular plants, we conducted a number of analyses: exclusion of gymnosperms and pteridophytes (‘angiosperms only’), exclusion of individual traits, and comparison to a gap-filled dataset representing about 15% of extant vascular plant species worldwide. The trait exclusion tests excluded the following individual traits, one at a time: LMA , N_{mass} and SSD , because analyses indicated that LMA and N_{mass} , although providing information on different aspects of leaf function (see Methods), are both part of the leaf economic spectrum^{7,10}, and SSD and plant height both reflect plant size when woody and herbaceous plants are considered together. To test if the results of the multivariate analysis presented in Fig. 2 were representative of vascular plants, we constructed a gap-filled dataset based on those species that entered the global dataset via the TRY Plant Trait Database. We extracted 328,057 individual plant-level trait observations, which provide a substantial number of additional data not used in the main analysis. We applied the data selection process as described above (section: Error detection and data quality control). The resulting dataset contained 78% missing entries (gaps), which were filled by Bayesian hierarchical probabilistic matrix factorization (BHPMF)^{253–255}. The gap-filled dataset was then used to calculate species mean trait values, resulting in a gap-filled dataset for 45,507 species. To quantitatively compare the results of the PCA presented in Fig. 2 and Extended Data Table 1 with those of the angiosperms-only and the gap-filled datasets, we applied a Procrustes test²⁵⁶ using the ‘procrustes’ and ‘protest’ functions in R package ‘vegan’²⁵⁷. Function protest tests the non-randomness between two configurations. Significant results (for example significance < 0.05) indicate that the shapes of two datasets are non-random to each other, but similar.

Kernel density estimation. To estimate the occurrence probability of given combinations of trait values in a two-dimensional space defined by PC axes 1 and 2 (Fig. 2), and bivariate trait combinations (Extended Data Fig. 4), we used two-dimensional kernel density estimation²⁵⁸. Because results depend on the choice of the bandwidth used for the smoothing kernel, we used unconstrained bandwidth selectors²⁵⁹. To visualize the occurrence probability of a given trait combination in the PCA space as well as for all possible bivariate trait combinations, we constructed contour plots from two-dimensional kernel density distributions. The colour gradient and contour lines in Fig. 2 and Extended Data Fig. 4 correspond to the 0.5, 0.95 and 0.99 quantiles of the respective probability distribution, thus highlighting the regions of highest and lowest trait occurrence probability. For kernel density estimation we used the ‘kde’ function and for optimal bandwidth selection carried out for each trait combination separately, we used the SAMSE pilot bandwidth selector²⁶⁰, both implemented in the R-package ‘ks’²⁵⁹. The R script used is provided at ftp://pbil.univ-lyon1.fr/pub/datasets/drays/Diaz_Nature/. **Code availability.** The R scripts used to generate the hypervolumes (Fig. 2) and kernel density analyses associated to Fig. 2 and Extended Data Fig. 4 are available at ftp://pbil.univ-lyon1.fr/pub/datasets/drays/Diaz_Nature/.

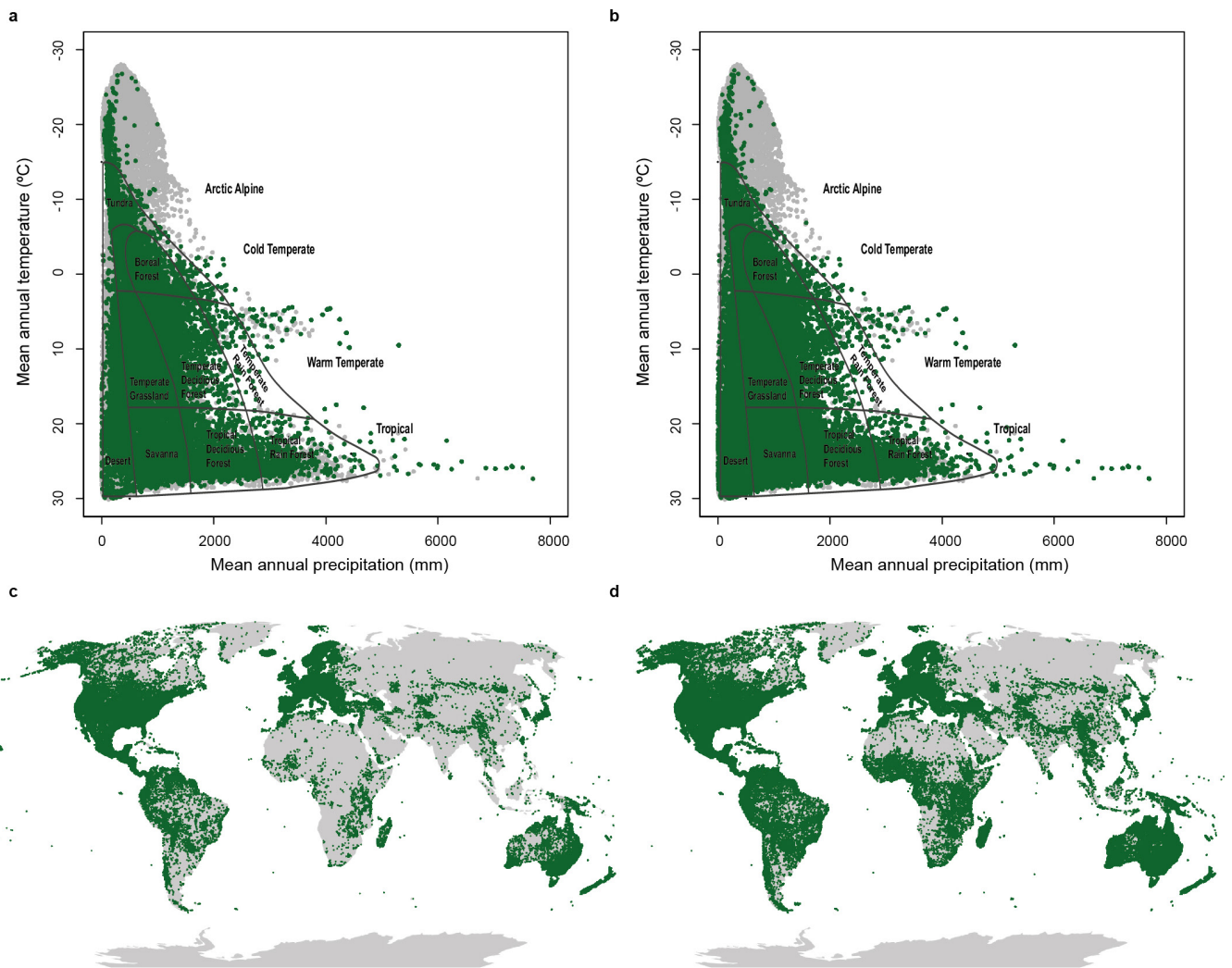
50. Gaudet, C. L. & Keddy, P. A. A comparative approach to predicting competitive ability from plant traits. *Nature* **334**, 242–243 (1988).
51. Westoby, M. A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant Soil* **199**, 213–227 (1998).
52. Muller-Landau, H. C., Wright, S. J., Calderón, O., Condit, R. & Hubbell, S. P. Interspecific variation in primary seed dispersal in a tropical forest. *J. Ecol.* **96**, 653–667 (2008).
53. Moles, A. T. et al. Global patterns in plant height. *J. Ecol.* **97**, 923–932 (2009).
54. Niklas, K. J. Influence of tissues density-specific mechanical-properties on the scaling of plant height. *Ann. Bot.* **72**, 173–179 (1993).
55. Donoghue, M. J. Key innovations, convergence, and success: macroevolutionary lessons from plant phylogeny. *Paleobiology* **31**, 77–93 (2005).
56. Preston, K. A., Cornwell, W. K. & Denoyer, J. L. Wood density and vessel traits as distinct correlates of ecological strategy in 51 California coast range angiosperms. *New Phytol.* **170**, 807–818 (2006).
57. Wright, S. J. et al. Functional traits and the growth-mortality trade-off in tropical trees. *Ecology* **91**, 3664–3674 (2010).

58. Givnish, T. J. & Vermeij, G. J. Sizes and shapes of liane leaves. *Am. Nat.* **110**, 743–778 (1976).
59. Givnish, T. J. Comparative studies of leaf form: assessing the relative roles of selective pressures and phylogenetic constraints. *New Phytol.* **106**, 131–160 (1987).
60. Farquhar, G. D., Buckley, T. N. & Miller, J. M. Optimal stomatal control in relation to leaf area and nitrogen content. *Silva Fennica* **36**, 625–637 (2002).
61. Ackerly, D. D. & Donoghue, M. J. Leaf size, sapling allometry, and Corner's rules: phylogeny and correlated evolution in maples (*Acer*). *Am. Nat.* **152**, 767–791 (1998).
62. Niklas, K. J. *Plant Allometry: the Scaling of Form and Process* (Univ. of Chicago Press, 1994).
63. Cornelissen, J. H. C. A triangular relationship between leaf size and seed size among woody species: allometry, ontogeny, ecology and taxonomy. *Oecologia* **118**, 248–255 (1999).
64. Poorter, H., Niinemets, U., Poorter, L., Wright, I. J. & Villar, R. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytol.* **182**, 565–588 (2009).
65. Bocanegra-González, K. T., Fernández-Mendez, F. & Galvis-Jiménez, J. F. Grupos funcionales de árboles en bosques secundarios de la región Bajo Calima (Buenaventura, Colombia). *Bol. Cient. Mus. Hist. Nat.* **19**, 17–40 (2015).
66. Thompson, K., Band, S. R. & Hodgson, J. G. Seed size and shape predict persistence in soil. *Funct. Ecol.* **7**, 236–241 (1993).
67. Moles, A. T. & Westoby, M. Seed size and plant strategy across the whole life cycle. *Oikos* **113**, 91–105 (2006).
68. Schleicher, A., Biedermann, R. & Kleyer, M. Dispersal traits determine plant response to habitat connectivity in an urban landscape. *Landscape Ecol.* **26**, 529–540 (2011).
69. Leishman, M. R. & Westoby, M. Hypotheses on seed size: tests using the semiarid flora of western New South Wales, Australia. *Am. Nat.* **143**, 890–906 (1994).
70. Muller-Landau, H. C. The tolerance-fecundity trade-off and the maintenance of diversity in seed size. *Proc. Natl Acad. Sci. USA* **107**, 4242–4247 (2010).
71. Thomson, F. J., Moles, A. T., Auld, T. D. & Kingsford, R. T. Seed dispersal distance is more strongly correlated with plant height than with seed mass. *J. Ecol.* **99**, 1299–1307 (2011).
72. Ben-Hur, E., Fragman-Sapir, O., Hadas, R., Singer, A. & Kadmon, R. Functional trade-offs increase species diversity in experimental plant communities. *Ecol. Lett.* **15**, 1276–1282 (2012).
73. Schulze, E.-D. & Schulze, I. Distribution and control of photosynthetic pathways in plants growing in the Namib Desert with special regard to *Welwitschia mirabilis*. *Madoqua* **9**, 5–13 (1976).
74. Schulze, E. D., Eller, B. M., Thomas, D. A., Willert, D. J. V. & Brinckmann, E. Leaf temperatures and energy balance of *Welwitschia mirabilis* in its natural habitat. *Oecologia* **44**, 258–262 (1980).
75. Zheng, W. *Silva Sinica: Volume 1–4* (China Forestry Publishing House, Beijing, 1983).
76. Shipley, B. & Parent, M. Germination responses of 64 wetland species in relation to seed size, minimum time to reproduction and seedling relative growth rate. *Funct. Ecol.* **5**, 111–118 (1991).
77. Fitter, A. H. & Peat, H. J. The ecological flora database. *J. Ecol.* **82**, 415–425 (1994).
78. Shipley, B. Structured interspecific determinants of specific leaf area in 34 species of herbaceous angiosperms. *Funct. Ecol.* **9**, 312–319 (1995).
79. Cornelissen, J. H. C. An experimental comparison of leaf decomposition rates in a wide range of temperate plant species and types. *J. Ecol.* **84**, 573–582 (1996).
80. Cornelissen, J. H. C., Diez, P. C. & Hunt, R. Seedling growth, allocation and leaf attributes in a wide range of woody plant species and types. *J. Ecol.* **84**, 755–765 (1996).
81. Atkin, O. K., Westbeek, M., Cambridge, M. L., Lambers, H. & Pons, T. L. Leaf respiration in light and darkness. *Plant Physiol.* **113**, 961–965 (1997).
82. Gillison, A. N. & Carpenter, G. A generic plant functional attribute set and grammar for dynamic vegetation description and analysis. *Funct. Ecol.* **11**, 775–783 (1997).
83. Castro-Diez, P., Puyravaud, J. P., Cornelissen, J. H. C. & Villar-Salvador, P. Stem anatomy and relative growth rate in seedlings of a wide range of woody plant species and types. *Oecologia* **116**, 57–66 (1998).
84. Atkin, O. K., Schortemeyer, M., McFarlane, N. & Evans, J. R. The response of fast- and slow-growing *Acacia* species to elevated atmospheric CO₂: an analysis of the underlying components of relative growth rate. *Oecologia* **120**, 544–554 (1999).
85. Bahn, M. et al. in *Land Use Changes in European Mountain Ecosystems: ECOMONT Concepts and Results* (eds Cernusca, A., Tappeiner, U. & Bayfield, N.) pp. 247–255 (Blackwell Wissenschaft, 1999).
86. Hickler, T. *Plant functional types and community characteristics along environmental gradients on Öland's Great Alvar (Sweden)*. MSc thesis, Univ. of Lund (1999).
87. Medlyn, B. E. et al. Effects of elevated [CO₂] on photosynthesis in European forest species: a meta-analysis of model parameters. *Plant Cell Environ.* **22**, 1475–1495 (1999).
88. Meziane, D. & Shipley, B. Interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability. *Plant Cell Environ.* **22**, 447–459 (1999).
89. Pyankov, V. I., Kondratchuk, A. V. & Shipley, B. Leaf structure and specific leaf mass: the alpine desert plants of the Eastern Pamirs, Tadjikistan. *New Phytol.* **143**, 131–142 (1999).
90. Carswell, F. E. et al. Photosynthetic capacity in a central Amazonian rain forest. *Tree Physiol.* **20**, 179–186 (2000).
91. Fonseca, C. R., Overton, J. M., Collins, B. & Westoby, M. Shifts in trait-combinations along rainfall and phosphorus gradients. *J. Ecol.* **88**, 964–977 (2000).
92. Shipley, B. & Lechowicz, M. J. The functional co-ordination of leaf morphology, nitrogen concentration, and gas exchange in 40 wetland species. *Ecoscience* **7**, 183–194 (2000).
93. Wilson, K. B., Baldocchi, D. D. & Hanson, P. J. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. *Tree Physiol.* **20**, 565–578 (2000).
94. Koike, F. Plant traits as predictors of woody species dominance in climax forest communities. *J. Veg. Sci.* **12**, 327–336 (2001).
95. Niinemets, Ü. Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology* **82**, 453–469 (2001).
96. Bond-Lamberty, B., Wang, C. & Gower, S. T. Aboveground and belowground biomass and sapwood area allometric equations for six boreal tree species of northern Manitoba. *Can. J. For. Res.* **32**, 1441–1450 (2002).
97. Bond-Lamberty, B., Wang, C., Gower, S. T. & Norman, J. Leaf area dynamics of a boreal black spruce fire chronosequence. *Tree Physiol.* **22**, 993–1001 (2002).
98. Meir, P. et al. Acclimation of photosynthetic capacity to irradiance in tree canopies in relation to leaf nitrogen concentration and leaf mass per unit area. *Plant Cell Environ.* **25**, 343–357 (2002).
99. Shipley, B. Trade-offs between net assimilation rate and specific leaf area in determining relative growth rate: relationship with daily irradiance. *Funct. Ecol.* **16**, 682–689 (2002).
100. Shipley, B. & Vu, T. T. Dry matter content as a measure of dry matter concentration in plants and their parts. *New Phytol.* **153**, 359–364 (2002).
101. Vendramini, F. et al. Leaf traits as indicators of resource-use strategy in floras with succulent species. *New Phytol.* **154**, 147–157 (2002).
102. Cornelissen, J. H. C. et al. Functional traits of woody plants: correspondence of species rankings between field adults and laboratory-grown seedlings? *J. Veg. Sci.* **14**, 311–322 (2003).
103. Loveys, B. R. et al. Thermal acclimation of leaf and root respiration: an investigation comparing inherently fast- and slow-growing plant species. *Glob. Change Biol.* **9**, 895–910 (2003).
104. Mencuccini, M. The ecological significance of long-distance water transport: short-term regulation, long-term acclimation and the hydraulic costs of stature across plant life forms. *Plant Cell Environ.* **26**, 163–182 (2003).
105. Ogaya, R. & Peñuelas, J. Comparative field study of *Quercus ilex* and *Phillyrea latifolia*: photosynthetic response to experimental drought conditions. *Environ. Exp. Bot.* **50**, 137–148 (2003).
106. Pillar, V. D. & Sosinski, E. E. An improved method for searching plant functional types by numerical analysis. *J. Veg. Sci.* **14**, 323–332 (2003).
107. Quedsted, H. M. et al. Decomposition of sub-arctic plants with differing nitrogen economies: a functional role for hemiparasites. *Ecology* **84**, 3209–3221 (2003).
108. Sack, L., Cowan, P. D., Jaikumar, N. & Holbrook, N. M. The 'hydrology' of leaves: co-ordination of structure and function in temperate woody species. *Plant Cell Environ.* **26**, 1343–1356 (2003).
109. Xu, L. & Baldocchi, D. D. Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (*Quercus douglasii*) under prolonged summer drought and high temperature. *Tree Physiol.* **23**, 865–877 (2003).
110. Adler, P. B., Milchunas, D. G., Lauenroth, W. K., Sala, O. E. & Burke, I. C. Functional traits of graminoids in semi-arid steppes: a test of grazing histories. *J. Appl. Ecol.* **41**, 653–663 (2004).
111. Cornelissen, J. H. C. et al. Leaf digestibility and litter decomposability are related in a wide range of subarctic plant species and types. *Funct. Ecol.* **18**, 779–786 (2004).
112. Fagúndez, J. & Izco, J. Seed morphology of *Calluna salisb.* (Ericaceae). *Acta Botánica Malacitana* **29**, 215–220 (2004).
113. Givnish, T. J., Montgomery, R. A. & Goldstein, G. Adaptive radiation of photosynthetic physiology in the Hawaiian lobeliads: light regimes, static light responses, and whole-plant compensation points. *Am. J. Bot.* **91**, 228–246 (2004).
114. Hill, M. O., Preston, C. D. & Roy, D. B. *PLANTATT - Attributes of British and Irish Plants: Status, Size, Life History, Geography and Habitats* <http://nora.nerc.ac.uk/9535/1/PLANTATT.pdf> (Centre for Ecology & Hydrology, NERC, 2004).
115. Kühn, I., Durka, W. & Klotz, S. BioFlor: a new plant-trait database as a tool for plant invasion ecology. *Divers. Distrib.* **10**, 363–365 (2004).
116. Sack, L. Responses of temperate woody seedlings to shade and drought: do trade-offs limit potential niche differentiation? *Oikos* **107**, 110–127 (2004).
117. Von Holle, B. & Simberloff, D. Testing Fox's assembly rule: does plant invasion depend on recipient community structure? *Oikos* **105**, 551–563 (2004).
118. Bakker, C., Rodenburg, J. & Van Bodegom, P. M. Effects of Ca- and Fe-rich seepage on P availability and plant performance in calcareous dune soils. *Plant Soil* **275**, 111–122 (2005).
119. Craine, J. M., Lee, W. G., Bond, W. J., Williams, R. J. & Johnson, L. C. Environmental constraints on a global relationship among leaf and root traits of grasses. *Ecology* **86**, 12–19 (2005).
120. Gachet, S., Vêla, E. & Taton, T. BASECO: a floristic and ecological database of Mediterranean French flora. *Biodivers. Conserv.* **14**, 1023–1034 (2005).

121. Han, W., Fang, J., Guo, D. & Zhang, Y. Leaf nitrogen and phosphorus stoichiometry across 753 terrestrial plant species in China. *New Phytol.* **168**, 377–385 (2005).
122. Kirkup, D., Malcolm, P., Christian, G. & Paton, A. Towards a digital African flora. *Taxon* **54**, 457–466 (2005).
123. Louault, F., Pillar, V. D., Aufrere, J., Garnier, E. & Soussana, J. F. Plant traits and functional types in response to reduced disturbance in a semi-natural grassland. *J. Veg. Sci.* **16**, 151–160 (2005).
124. Moles, A. T. *et al.* A brief history of seed size. *Science* **307**, 576–580 (2005).
125. Nakahashi, C. D., Frole, K. & Sack, L. Bacterial leaf nodule symbiosis in *Ardisia* (Myrsinaceae): does it contribute to seedling growth capacity? *Plant Biol.* **7**, 495–500 (2005).
126. Peco, B., de Pablos, I., Traba, J. & Levassor, C. The effect of grazing abandonment on species composition and functional traits: the case of dehesa grasslands. *Basic Appl. Ecol.* **6**, 175–183 (2005).
127. Rentería, L., Jaramillo, V. J., Martínez-Yrizar, A. & Pérez-Jiménez, A. Nitrogen and phosphorus resorption in trees of a Mexican tropical dry forest. *Trees* **19**, 431–441 (2005).
128. Sack, L., Tyree, M. T. & Holbrook, N. M. Leaf hydraulic architecture correlates with regeneration irradiance in tropical rainforest trees. *New Phytol.* **167**, 403–413 (2005).
129. Schweingruber, F. H. & Landolt, W. The xylem database. <http://www.wsl.ch/dendropro/xylemldb/> Swiss Federal Research Institute WSL, updated 2014.
130. Sheremet'ev, S. N. Herbs on the soil moisture gradient (water relations and the structural-functional organization). *KMK, Moscow* **271** (2005).
131. Vile, D. *Significations Fonctionnelle et Ecologique des Traits des Espèces Vegetales: Exemple dans une Succession Post-Culturelle Méditerranéenne et Généralisations*. PhD thesis, Univ. Montpellier 2 (2005).
132. Bakker, C., Van Bodegom, P. M., Nelissen, H. J. M., Ernst, W. H. O. & Aerts, R. Plant responses to rising water tables and nutrient management in calcareous dune slacks. *Plant Ecol.* **185**, 19–28 (2006).
133. Cavender-Bares, J., Keen, A. & Miles, B. Phylogenetic structure of Floridian plant communities depends on taxonomic and spatial scale. *Ecology* **87** (Suppl.), S109–S122 (2006).
134. Kazakou, E., Vile, D., Shipley, B., Gallet, C. & Garnier, E. Co-variations in litter decomposition, leaf traits and plant growth in species from a Mediterranean old-field succession. *Funct. Ecol.* **20**, 21–30 (2006).
135. Kerkhoff, A. J. & Enquist, B. J. Ecosystem allometry: the scaling of nutrient stocks and primary productivity across plant communities. *Ecol. Lett.* **9**, 419–427 (2006).
136. Poorter, L. & Bongers, F. Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology* **87**, 1733–1743 (2006).
137. Sack, L. & Frole, K. Leaf structural diversity is related to hydraulic capacity in tropical rain forest trees. *Ecology* **87**, 483–491 (2006).
138. Sack, L., Melcher, P. J., Liu, W. H., Middleton, E. & Pardee, T. How strong is intracanopy leaf plasticity in temperate deciduous trees? *Am. J. Bot.* **93**, 829–839 (2006).
139. Campbell, C. *et al.* Acclimation of photosynthesis and respiration is asynchronous in response to changes in temperature regardless of plant functional group. *New Phytol.* **176**, 375–389 (2007).
140. Cavender-Bares, J., Sack, L. & Savage, J. Atmospheric and soil drought reduce nocturnal conductance in live oaks. *Tree Physiol.* **27**, 611–620 (2007).
141. Cornwell, W. K., Bhaskar, R., Sack, L. & Cordell, S. Adjustment of structure and function of Hawaiian *Metrosideros polymorpha* at high vs. low precipitation. *Funct. Ecol.* **21**, 1063–1071 (2007).
142. Craven, D. *et al.* Between and within-site comparisons of structural and physiological characteristics and foliar nutrient content of 14 tree species at a wet, fertile site and a dry, infertile site in Panama. *For. Ecol. Manage.* **238**, 335–346 (2007).
143. Domingues, T. F., Martinelli, L. A. & Ehleringer, J. R. Ecophysiological traits of plant functional groups in forest and pasture ecosystems from eastern Amazonia, Brazil. *Plant Ecol.* **193**, 101–112 (2007).
144. Garnier, E. *et al.* Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: a standardized methodology and lessons from an application to 11 European sites. *Ann. Bot.* **99**, 967–985 (2007).
145. Laverigne, S. & Molofsky, J. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proc. Natl Acad. Sci. USA* **104**, 3883–3888 (2007).
146. Meir, P., Levy, P. E., Grace, J. & Jarvis, P. G. Photosynthetic parameters from two contrasting woody vegetation types in West Africa. *Plant Ecol.* **192**, 277–287 (2007).
147. Müller, S. C., Overbeck, G. E., Pfadenhauer, J. & Pillar, V. D. Plant functional types of woody species related to fire disturbance in forest–grassland ecotones. *Plant Ecol.* **189**, 1–14 (2007).
148. Pierce, S., Ceriani, R. M., De Andreis, R., Luzzaro, A. & Cerabolini, B. The leaf economics spectrum of Poaceae reflects variation in survival strategies. *Plant Biosyst.* **141**, 337–343 (2007).
149. Pierce, S., Luzzaro, A., Caccianiga, M., Ceriani, R. M. & Cerabolini, B. Disturbance is the principal α -scale filter determining niche differentiation, coexistence and biodiversity in an alpine community. *J. Ecol.* **95**, 698–706 (2007).
150. Price, C. A. & Enquist, B. J. Scaling mass and morphology in leaves: an extension of the WBE model. *Ecology* **88**, 1132–1141 (2007).
151. Scherer-Lorenzen, M., Schulze, E.-D., Don, A., Schumacher, J. & Weller, E. Exploring the functional significance of forest diversity: a new long-term experiment with temperate tree species (BIOTREE). *Perspect. Plant Ecol. Evol. Syst.* **9**, 53–70 (2007).
152. Schurr, F. M. *et al.* Colonization and persistence ability explain the extent to which plant species fill their potential range. *Glob. Ecol. Biogeogr.* **16**, 449–459 (2007).
153. Swaine, E. K. *Ecological and Evolutionary Drivers of Plant Community Assembly in a Bornean Rain Forest*. PhD thesis, Univ. of Aberdeen (2007).
154. Vaieretti, M. V., Díaz, S., Vile, D. & Garnier, E. Two measurement methods of leaf dry matter content produce similar results in a broad range of species. *Ann. Bot.* **99**, 955–958 (2007).
155. Coomes, D. A., Heathcote, S., Godfrey, E. R., Shepherd, J. J. & Sack, L. Scaling of xylem vessels and veins within the leaves of oak species. *Biol. Lett.* **4**, 302–306 (2008).
156. Cornwell, W. K. *et al.* Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecol. Lett.* **11**, 1065–1071 (2008).
157. Hoof, J., Sack, L., Webb, D. T. & Nilsen, E. T. Contrasting structure and function of pubescent and glabrous varieties of Hawaiian *Metrosideros polymorpha* (Myrtaceae) at high elevation. *Biotropica* **40**, 113–118 (2008).
158. Kleyer, M. *et al.* The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *J. Ecol.* **96**, 1266–1274 (2008).
159. Kraft, N. J. B., Valencia, R. & Ackerly, D. D. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* **322**, 580–582 (2008).
160. Kurokawa, H. & Nakashizuka, T. Leaf herbivory and decomposability in a Malaysian tropical rain forest. *Ecology* **89**, 2645–2656 (2008).
161. Quero, J. L. *et al.* Relating leaf photosynthetic rate to whole-plant growth: drought and shade effects on seedlings of four *Quercus* species. *Funct. Plant Biol.* **35**, 725–737 (2008).
162. Reich, P. B. *et al.* Scaling of respiration to nitrogen in leaves, stems and roots of higher land plants. *Ecol. Lett.* **11**, 793–801 (2008).
163. Royal Botanical Gardens Kew Seed Information Database (SID). Version 7.1. <http://data.kew.org/sid/> (2015).
164. Scoffoni, C., Pou, A., Aasamaa, K. & Sack, L. The rapid light response of leaf hydraulic conductance: new evidence from two experimental methods. *Plant Cell Environ.* **31**, 1803–1812 (2008).
165. Shiodera, S., Rahajoe, J. S. & Kohyama, T. Variation in longevity and traits of leaves among co-occurring understorey plants in a tropical montane forest. *J. Trop. Ecol.* **24**, 121–133 (2008).
166. van Bodegom, P. M., Sorrell, B. K., Oosthoek, A., Bakker, C. & Aerts, R. Separating the effects of partial submergence and soil oxygen demand on plant physiology. *Ecology* **89**, 193–204 (2008).
167. Bragazza, L. Conservation priority of Italian Alpine habitats: a floristic approach based on potential distribution of vascular plant species. *Biodivers. Conserv.* **18**, 2823–2835 (2009).
168. Ciocarlan, V. *The Illustrated Flora of Romania. Pteridophyta et Spermatopyta*. (Editura Ceres Bucharest, 2009).
169. Craine, J. M. *et al.* Global patterns of foliar nitrogen isotopes and their relationships with climate, mycorrhizal fungi, foliar nutrient concentrations, and nitrogen availability. *New Phytol.* **183**, 980–992 (2009).
170. Dunbar, S., Sporck, M. & Sack, L. Leaf trait diversification and design in seven rare taxa of the Hawaiian *Plantago* radiation. *Int. J. Plant Sci.* **170**, 61–75 (2009).
171. Fyllas, N. M. *et al.* Basin-wide variations in foliar properties of Amazonian forest: phylogeny, soils and climate. *Biogeosciences* **6**, 2677–2708 (2009).
172. Green, W. USDA PLANTS Compilation, Version 1. <http://bricol.net/downloads/data/PLANTSdatabase/> National Plant Data Center Baton Rouge (2009).
173. Kattge, J., Knorr, W., Raddatz, T. & Wirth, C. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Glob. Change Biol.* **15**, 976–991 (2009).
174. Malhado, A. C. M. *et al.* Spatial trends in leaf size of Amazonian rainforest trees. *Biogeosciences* **6**, 1563–1576 (2009).
175. Manning, P., Houston, K. & Evans, T. Shifts in seed size across experimental nitrogen enrichment and plant density gradients. *Basic Appl. Ecol.* **10**, 300–308 (2009).
176. Moretti, M. & Legg, C. Combining plant and animal traits to assess community functional responses to disturbance. *Ecography* **32**, 299–309 (2009).
177. Paula, S. *et al.* Fire-related traits for plant species of the Mediterranean Basin: Ecological Archives E090-094. *Ecology* **90**, 1420 (2009).
178. Poorter, L. Leaf traits show different relationships with shade tolerance in moist versus dry tropical forests. *New Phytol.* **181**, 890–900 (2009).
179. Reich, P. B., Oleksyn, J. & Wright, I. J. Leaf phosphorus influences the photosynthesis–nitrogen relation: a cross-biome analysis of 314 species. *Oecologia* **160**, 207–212 (2009).
180. Tecco, P. A., Diaz, S., Cabido, M. & Urcelay, C. Functional traits of alien plants across contrasting climatic and land-use regimes: do aliens join the locals or try harder than them? *J. Ecol.* **98**, 17–27 (2009).
181. Wirth, C. & Lichstein, J. W. *The Imprint of Species Turnover on Old-Growth Forest Carbon Balances-Insights from a Trait-Based Model of Forest Dynamics in Old-growth Forests - Function, Fate and Value* 81–113 (eds. Wirth, C., Gleixner, G. & Heimann, M.) (Springer-Verlag, 2009).
182. Cerabolini, B. E. L. *et al.* Can CSR classification be generally applied outside Britain? *Plant Ecol.* **210**, 253–261 (2010).
183. Domingues, T. F. *et al.* Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant Cell Environ.* **33**, 959–980 (2010).

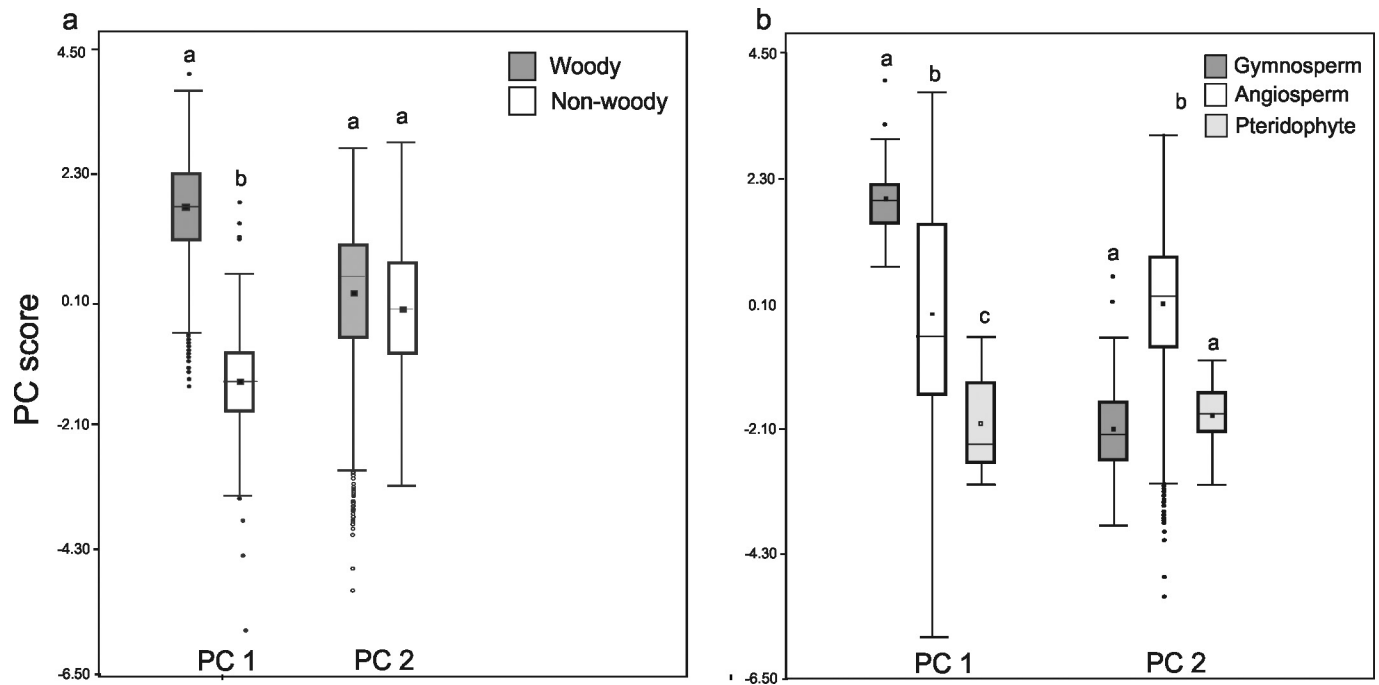
184. Hao, G. Y., Sack, L., Wang, A. Y., Cao, K. F. & Goldstein, G. Differentiation of leaf water flux and drought tolerance traits in hemiepiphytic and non-hemiepiphytic *Ficus* tree species. *Funct. Ecol.* **24**, 731–740 (2010).
185. Messier, J., McGill, B. J. & Lechowicz, M. J. How do traits vary across ecological scales? A case for trait-based ecology. *Ecol. Lett.* **13**, 838–848 (2010).
186. Ordoñez, J. C. *et al.* Plant strategies in relation to resource supply in mesic to wet environments: does theory mirror nature? *Am. Nat.* **175**, 225–239 (2010).
187. Peñuelas, J. *et al.* Faster returns on 'leaf economics' and different biogeochemical niche in invasive compared with native plant species. *Glob. Change Biol.* **16**, 2171–2185 (2010).
188. Powers, J. S. & Tiffin, P. Plant functional type classifications in tropical dry forests in Costa Rica: leaf habit versus taxonomic approaches. *Funct. Ecol.* **24**, 927–936 (2010).
189. Willis, C. G. *et al.* Phylogenetic community structure in Minnesota oak savanna is influenced by spatial extent and environmental variation. *Ecography* **33**, 565–577 (2010).
190. Blonder, B., Violle, C., Bentley, L. P. & Enquist, B. J. Venation networks and the origin of the leaf economics spectrum. *Ecol. Lett.* **14**, 91–100 (2011).
191. Brown, K. A. *et al.* Assessing natural resource use by forest-reliant communities in Madagascar using functional diversity and functional redundancy metrics. *PLoS One* **6**, e24107 (2011).
192. Butterfield, B. J. & Briggs, J. M. Regeneration niche differentiates functional strategies of desert woody plant species. *Oecologia* **165**, 477–487 (2011).
193. Campetella, G. *et al.* Patterns of plant trait–environment relationships along a forest succession chronosequence. *Agric. Ecosyst. Environ.* **145**, 38–48 (2011).
194. Craine, J. M. *et al.* Functional consequences of climate change-induced plant species loss in a tallgrass prairie. *Oecologia* **165**, 1109–1117 (2011).
195. de Araujo, A. C. *et al.* LBA-ECO CD-02 C and N Isotopes in Leaves and Atmospheric CO₂, Amazonas, Brazil. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA <http://dx.doi.org/10.3334/ORNLDAAAC/1097> (2012).
196. Laughlin, D. C., Fule, P. Z., Huffman, D. W., Crouse, J. & Laliberte, E. Climatic constraints on trait-based forest assembly. *J. Ecol.* **99**, 1489–1499 (2011).
197. Milla, R. & Reich, P. B. Multi-trait interactions, not phylogeny, fine-tune leaf size reduction with increasing altitude. *Ann. Bot.* **107**, 455–465 (2011).
198. Onoda, Y. *et al.* Global patterns of leaf mechanical properties. *Ecol. Lett.* **14**, 301–312 (2011).
199. Prentice, I. C. *et al.* Evidence of a universal scaling relationship for leaf CO₂ drawdown along an aridity gradient. *New Phytol.* **190**, 169–180 (2011).
200. Sandel, B., Corbin, J. D. & Krupa, M. Using plant functional traits to guide restoration: a case study in California coastal grassland. *Ecosphere* **2**, art23 (2011).
201. Tucker, S. S., Craine, J. M. & Nippert, J. B. Physiological drought tolerance and the structuring of tallgrass prairie assemblages. *Ecosphere* **2**, art48 (2011).
202. Yguel, B. *et al.* Phytophagy on phylogenetically isolated trees: why hosts should escape their relatives. *Ecol. Lett.* **14**, 1117–1124 (2011).
203. Adriaenssens, S. *Dry Deposition and Canopy Exchange for Temperate Tree Species under High Nitrogen Deposition*. PhD thesis, Ghent Univ. (2012).
204. Beckmann, M., Hock, M., Bruehlheide, H. & Erfmeier, A. The role of UV-B radiation in the invasion of *Hieracium pilosella*—A comparison of German and New Zealand plants. *Environ. Exp. Bot.* **75**, 173–180 (2012).
205. Choat, B. *et al.* Global convergence in the vulnerability of forests to drought. *Nature* **491**, 752–755 (2012).
206. Craine, J., Towne, E. G., Ocheltree, T. & Nippert, J. Community traitscape of foliar nitrogen isotopes reveals N availability patterns in a tallgrass prairie. *Plant Soil* **356**, 395–403 (2012).
207. Frenette-Dussault, C., Shipley, B., Léger, J. F., Meziane, D. & Hingrat, Y. Functional structure of an arid steppe plant community reveals similarities with Grime's C-S-R theory. *J. Veg. Sci.* **23**, 208–222 (2012).
208. Gallagher, R. V. & Leishman, M. R. A global analysis of trait variation and evolution in climbing plants. *J. Biogeogr.* **39**, 1757–1771 (2012).
209. Guerin, G. R., Wen, H. & Lowe, A. J. Leaf morphology shift linked to climate change. *Biol. Lett.* **8**, 882–886 (2012).
210. Gutiérrez, A. G. & Huth, A. Successional stages of primary temperate rainforests of Chiloé Island, Chile. *Perspect. Plant Ecol. Evol. Syst.* **14**, 243–256 (2012).
211. Guy, A. L., Mischkolz, J. M. & Lamb, E. G. Limited effects of simulated acidic deposition on seedling survivorship and root morphology of endemic plant taxa of the Athabasca Sand Dunes in well-watered greenhouse trials. *Botany* **91**, 176–181 (2012).
212. Han, W. *et al.* Floral, climatic and soil pH controls on leaf ash content in China's terrestrial plants. *Glob. Ecol. Biogeogr.* **21**, 376–382 (2012).
213. Minden, V., Andratschke, S., Spalke, J., Timmermann, H. & Kleyer, M. Plant trait–environment relationships in salt marshes: Deviations from predictions by ecological concepts. *Perspect. Plant Ecol. Evol. Syst.* **14**, 183–192 (2012).
214. Minden, V. & Kleyer, M. Testing the effect–response framework: key response and effect traits determining above-ground biomass of salt marshes. *J. Veg. Sci.* **22**, 387–401 (2011).
215. Pierce, S., Brusa, G., Sartori, M. & Cerabolini, B. E. L. Combined use of leaf size and economics traits allows direct comparison of hydrophyte and terrestrial herbaceous adaptive strategies. *Ann. Bot.* **109**, 1047–1053 (2012).
216. Spasojevic, M. J. & Suding, K. N. Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *J. Ecol.* **100**, 652–661 (2012).
217. Vergutz, L., Manzoni, S., Porporato, A., Novais, R. F. & Jackson, R. B. A Global Database of Carbon and Nutrient Concentrations of Green and Senesced Leaves. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA <http://dx.doi.org/10.3334/ORNLDAAAC/1106> (2012).
218. Williams, M., Shimabokuro, Y. E. & Rastetter, E. B. LBA-ECO CD-09 Soil and Vegetation Characteristics, Tapajós National Forest, Brazil. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA <http://dx.doi.org/10.3334/ORNLDAAAC/1104> (2012).
219. Wright, J. P. & Sutton-Grier, A. Does the leaf economic spectrum hold within local species pools across varying environmental conditions? *Funct. Ecol.* **26**, 1390–1398 (2012).
220. Auger, S. & Shipley, B. Interspecific and intraspecific trait variation along short environmental gradients in an old-growth temperate forest. *J. Veg. Sci.* **24**, 419–428 (2013).
221. Blonder, B., Violle, C. & Enquist, B. J. Assessing the causes and scales of the leaf economics spectrum using venation networks in *Populus tremuloides*. *J. Ecol.* **101**, 981–989 (2013).
222. Byun, C., Blois, S. & Brissot, J. Plant functional group identity and diversity determine biotic resistance to invasion by an exotic grass. *J. Ecol.* **101**, 128–139 (2013).
223. Chen, Y., Han, W., Tang, L., Tang, Z. & Fang, J. Leaf nitrogen and phosphorus concentrations of woody plants differ in responses to climate, soil and plant growth form. *Ecography* **36**, 178–184 (2013).
224. Conti, G. & Díaz, S. Plant functional diversity and carbon storage—an empirical test in semi-arid forest ecosystems. *J. Ecol.* **101**, 18–28 (2013).
225. Conti, G., Enrico, L., Casanoves, F. & Díaz, S. Shrub biomass estimation in the semiarid Chaco forest: a contribution to the quantification of an underrated carbon stock. *Ann. For. Sci.* **70**, 515–524 (2013).
226. Demey, A. *et al.* Nutrient input from hemiparasitic litter favors plant species with a fast-growth strategy. *Plant Soil* **371**, 53–66 (2013).
227. Kichenin, E., Wardle, D. A., Peltzer, D. A., Morse, C. W. & Freschet, G. T. Contrasting effects of plant inter- and intraspecific variation on community-level trait measures along an environmental gradient. *Funct. Ecol.* **27**, 1254–1261 (2013).
228. Pahl, A. T., Kollmann, J., Mayer, A. & Haider, S. No evidence for local adaptation in an invasive alien plant: field and greenhouse experiments tracing a colonization sequence. *Ann. Bot.* **112**, 1921–1930 (2013).
229. Brando, P. M. *et al.* Abrupt increases in Amazonian tree mortality due to drought-fire interactions. *Proc. Natl Acad. Sci. USA* **111**, 6347–6352 (2014).
230. Everwand, G., Fry, E. L., Eggers, T. & Manning, P. Seasonal variation in the capacity for plant trait measures to predict grassland carbon and water fluxes. *Ecosystems* (N. Y.) **17**, 1095–1108 (2014).
231. Fry, E. L., Power, S. A. & Manning, P. Trait-based classification and manipulation of plant functional groups for biodiversity–ecosystem function experiments. *J. Veg. Sci.* **25**, 248–261 (2014).
232. Iida, Y. *et al.* Linking functional traits and demographic rates in a subtropical tree community: the importance of size dependency. *J. Ecol.* **102**, 641–650 (2014).
233. Blonder, B. *et al.* Testing models for the leaf economics spectrum with leaf and whole-plant traits in *Arabidopsis thaliana*. *AoB Plants* **7**, plv049 (2015).
234. Meng, T.-T. *et al.* Responses of leaf traits to climatic gradients: adaptive variation versus compositional shifts. *Biogeosciences Discuss.* **12**, 7093–7124 (2015).
235. Burrascano, S. *et al.* Environmental filtering of wild boar rooting activity on understorey composition and functional traits. *Community Ecol.* **16**, 244–253 (2015).
236. Albert, C. H. *et al.* Intraspecific functional variability: extent, structure and sources of variation. *J. Ecol.* **98**, 604–613 (2010).
237. Gillison, A. N. in *Vegetation Ecology* 2nd edn (eds. van der Maarel, E. & Franklin, J.) pp. 347–386 (Wiley Online Library, 2013).
238. Niinemets, Ü., Keenan, T. F. & Hallik, L. A worldwide analysis of within-canopy variations in leaf structural, chemical and physiological traits across plant functional types. *New Phytol.* **205**, 973–993 (2015).
239. Pérez-Harguindeguy, N. *et al.* New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* **61**, 167–234 (2013).
240. Page, C. N. *The ferns of Britain and Ireland* (Cambridge Univ. Press, 1997).
241. Lloyd, R. M. Spore morphology of the Hawaiian genus *Sadleria* (Blechnaceae). *Am. Fern J.* **66**, 1–7 (1976).
242. Conway, E. Spore production in bracken (*Pteridium aquilinum* (L.) Kuhn). *J. Ecol.* **45**, 273–284 (1957).
243. Stoor, A. M., Boudrie, M., Jérôme, C., Horn, K. & Bennert, H. W. *Diphysastrum oellgaardii* (Lycopodiaceae, Pteridophyta), a new lycopod species from Central Europe and France. *Feddes Repert.* **107**, 149–157 (1996).
244. Büntgen, U., Psomas, A. & Schweingruber, F. H. Introducing wood anatomical and dendrochronological aspects of herbaceous plants: applications of the Xylem Database to vegetation science. *J. Veg. Sci.* **25**, 967–977 (2014).
245. Blonder, B., Lamanna, C., Violle, C. & Enquist, B. J. The *n*-dimensional hypervolume. *Glob. Ecol. Biogeogr.* **23**, 595–609 (2014).
246. Manly, B. J. *Randomization, Bootstrap and Monte Carlo Methods in Biology* (Chapman & Hall, 1997).
247. Di Rienzo, J. A. *et al.* InfoStat version 2015. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. (Córdoba, 2015).
248. Jongman, R. H. G., ter Braak, C. J. F. & van Tongeren, O. F. R. *Data Analysis in Community and Landscape Ecology*. (Pudor, Wageningen, 1987).

249. Dray, S. On the number of principal components: a test of dimensionality based on measurements of similarity between matrices. *Comput. Stat. Data Anal.* **52**, 2228–2237 (2008).
250. Pinheiro, J., Bates, D., DebRoy, S. & Sarkar, D. nlme: Linear and Nonlinear Mixed Effects Models. “nlme” R package version 3.1-108. <http://CRAN.R-project.org/package=nlme> (2013).
251. Bates, D., Maechler, M. & Bolker, B. lme4: linear mixed-effects models using Eigen and R. R package version 0.999999-2. <http://lme4.r-forge.r-project.org/> (2013).
252. R Development Core Team. R: A language and environment for statistical computing. <https://www.r-project.org/> (R Foundation for Statistical Computing, 2014).
253. Shan, H. *et al.* Gap filling in the Plant Kingdom: trait prediction using Hierarchical Probabilistic Matrix Factorization. *Proceedings of the 29th International Conference on Machine Learning (ICML)* <http://arxiv.org/abs/1206.6439> (2012).
254. Fazayeli, F., Banerjee, A., Kattge, J., Schrod, F. & Reich, P. B. Uncertainty quantified matrix completion using Bayesian Hierarchical Matrix Factorization in *13th International Conference on Machine Learning and Applications (ICMLA)*, <http://dx.doi.org/10.1109/ICMLA.2014.56> (2014).
255. Schrod, F. *et al.* BHPMF – a hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography. *Glob. Ecol. Biogeogr.* **24**, 1510–1521 (2015).
256. Peres-Neto, P. R. & Jackson, D. A. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* **129**, 169–178 (2001).
257. Oksanen, J. *et al.* vegan: Community Ecology Package, version 2.0–7. <http://CRAN.R-project.org/package=vegan> (2008).
258. Wand, M. & Jones, M. *Kernel Smoothing* (Chapman and Hall, 1995).
259. Duong, T. ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *J. Stat. Softw.* **21**, 1–16 (2007).
260. Duong, T. & Hazelton, M. Plug-in bandwidth matrices for bivariate kernel density Estimation. *J. Nonparametr. Stat.* **15**, 17–30 (2003).
261. New, M., Hulme, M. & Jones, P. Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J. Clim.* **12**, 829–856 (1999).
262. Whittaker, R. J. *Communities and Ecosystems* (Macmillan, 1975).
263. Warton, D. I., Wright, I. J., Falster, D. S. & Westoby, M. Bivariate line-fitting methods for allometry. *Biol. Rev. Camb. Philos. Soc.* **81**, 259–291 (2006).



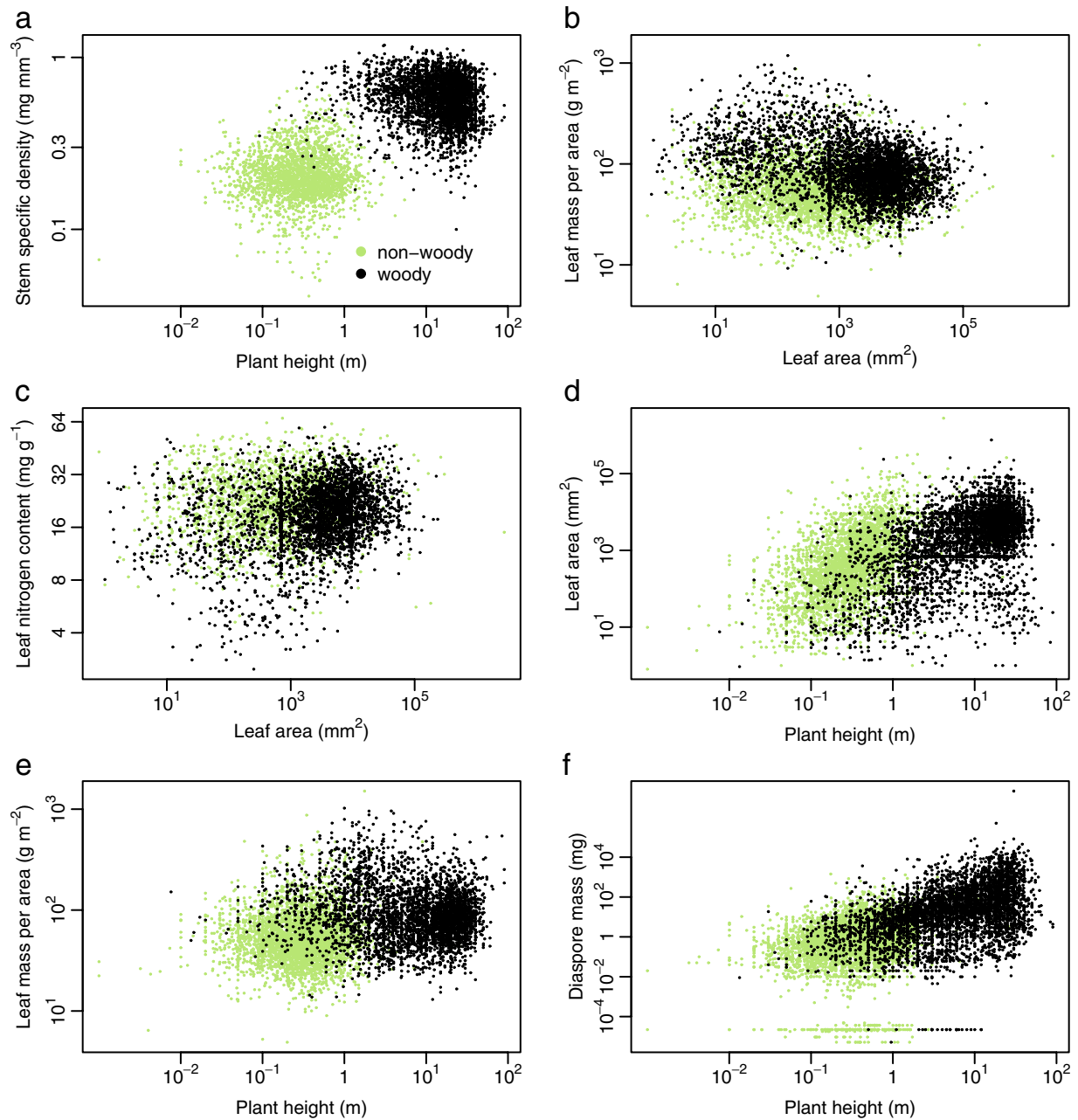
Extended Data Figure 1 | Climatic and geographical coverage of the dataset. **a–d**, Green points, occurrence according to GBIF (<http://www.gbif.org>) of species with information on all six traits (**a, c**) and at least one trait (**b, d**). Upper panels (**a, b**) show distribution in major climatic regions of the world; grey, MAP and MAT as in CRU0.5 degree climatology²⁶¹,

Biome classification according to ref. 262. Lower panels (**c, d**) show distribution in the global map (Robinson projection); grey, land surface. Maps based on the R package 'maps', accessed at The Comprehensive R Archive Network (<https://cran.r-project.org/web/packages/maps/index.html>).

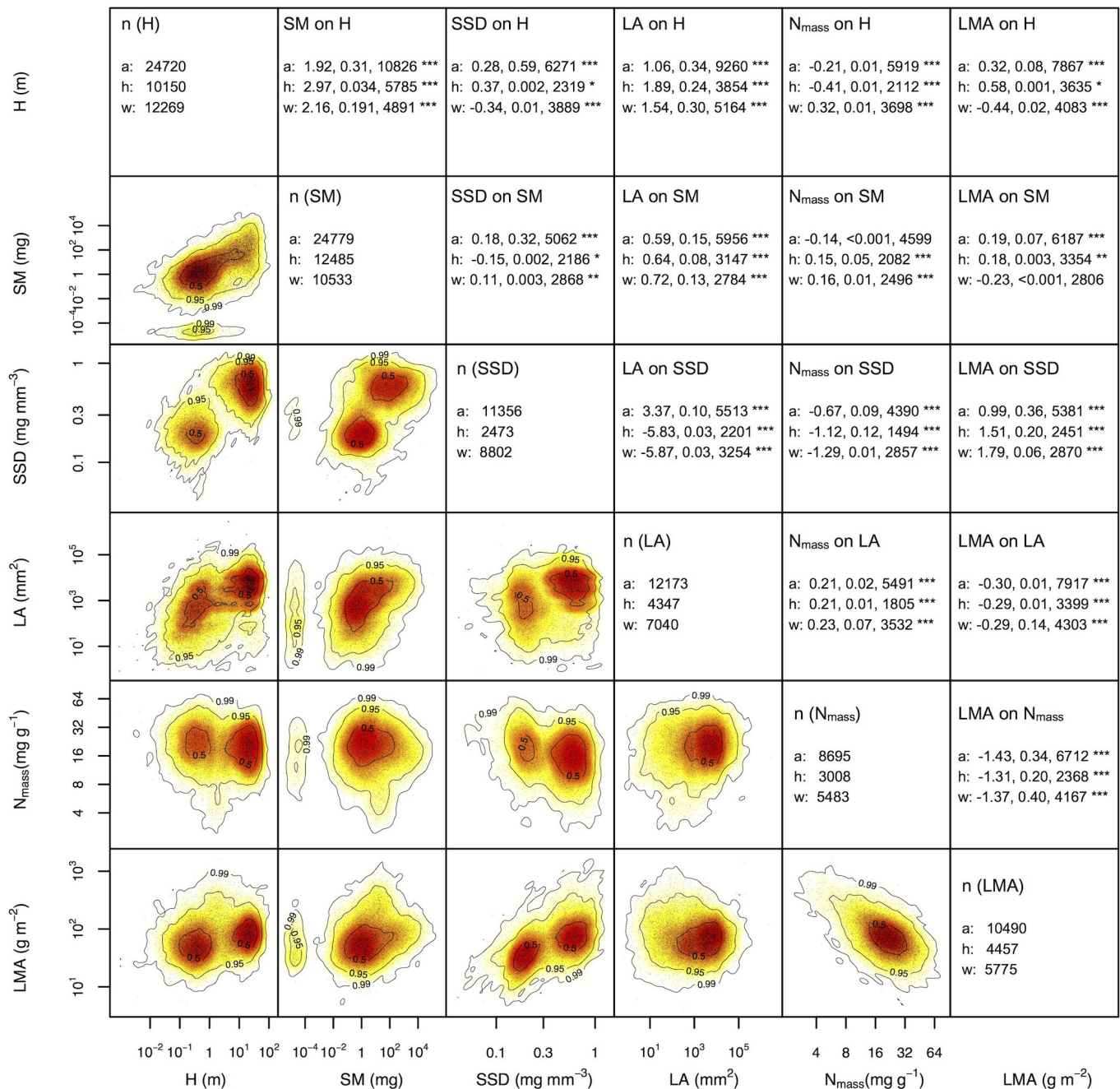


Extended Data Figure 2 | Tests of the distribution of growth-forms (a) and major taxa (b) in trait space. Woody and non-woody species differed significantly in their positions along PC1 but not along PC2. Angiosperms differed significantly from gymnosperms and pteridophytes in their positions along both axes; gymnosperms and pteridophytes differed in their position along PC1 but not along PC2 (see Methods for details

of PCA analysis and a posteriori tests). Whiskers denote ± 3 s.d. from mean; n woody = 1,001; n non-woody = 1,209; n angiosperms = 2,120; n gymnosperms = 80; n pteridophytes = 14). Horizontal bars and dots within boxes indicate mean and median, respectively. Means with the same letter are not significantly different (Fisher's least significant difference; $P > 0.01$).



Extended Data Figure 3 | Selected bivariate relationships underlying the global spectrum of plant form and function, showing herbaceous (green) and woody (black) species separately. See Extended Data Fig. 4 for standardized major axes statistics (slope, r^2 , sample size) of these and all other pairwise trait combinations.



Extended Data Figure 4 | Bivariate relationships between the six traits that underlie the global spectrum of plant form and function. The lower left portion of the matrix shows two-dimensional probability density distributions of bivariate trait relationships derived through kernel density estimation (see Methods). The colour gradient indicates regions of highest (red) to lowest (white) occurrence probability of trait combinations with contour lines indicating 0.5, 0.95 and 0.99 quantiles. The upper right portion contains standardized major axis (SMA)²⁶³ statistics (slope, r^2 , sample size n , and statistical significance, NS, $P > 0.05$; $*0.05 > P > 0.01$; $**0.01 > P > 0.001$; $***P < 0.001$) for the corresponding

relationships for all species (a), and for herbaceous (h) and woody species (w) separately. The diagonal displays the total sample sizes for each trait. For traits showing a strongly bimodal distribution, the all-species slope and correlation should be treated with caution. Pteridophytes show a discontinuous distribution in SM, but otherwise fall well within the general distribution of points; they represent less than 1% of the dataset, therefore including or excluding them does not significantly alter any of the relationships. SMAs were fitted using SMATR v.2 (<http://www.bio.mq.edu.au/ecology/SMATR/>).

Extended Data Table 1 | Principal component analyses (PCAs) of global plant trait data

	Main analysis		Angiosperms only		Excl. LMA		Excl. N _{mass}		Excl. SSD		Gap-filled	
	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2	PC1	PC2
Variation explained (%)	49	25	50	24	52	24	56	23	45	30	42	24
Eigenvalue	2.93	1.50	3.01	1.44	2.60	1.19	2.81	1.17	2.26	1.48	2.53	1.44
Significance of Procrustes test			0.0014								0.0014	
Variable loadings												
Leaf area (LA)	0.23	0.58	0.30	0.51	0.33	0.56	0.28	0.73	0.34	0.53	0.34	0.44
Leaf nitrogen per mass (N _{mass})	-0.25	0.57	-0.21	0.60	-0.17	0.76	-	-	-0.27	0.62	-0.21	0.65
Leaf mass per area (LMA)	0.40	-0.46	0.37	-0.49	-	-	0.36	-0.60	0.42	-0.52	0.36	-0.56
Plant height (H)	0.52	0.20	0.52	0.22	0.57	0.02	0.54	0.10	0.59	0.13	0.54	0.09
Diaspore mass (SM)	0.45	0.30	0.46	0.25	0.51	0.16	0.49	0.19	0.54	0.23	0.49	0.24
Stem specific density (SSD)	0.51	-0.09	0.50	-0.14	0.52	-0.30	0.51	-0.26	-	-	0.42	0.03

Eigenvalues and trait loadings of principal components (PC1 and PC2) in six different PCAs. Main analysis corresponds to the PCA performed on 2,214 species for which values of all six traits were available, and which is reported in the main text and expressed graphically in Fig. 2. The rest of the columns correspond to PCAs carried out on angiosperms only (2,120 species), on all taxa but excluding LMA, N_{mass} or SSD one at a time (2,214 in all cases), and on a gap-filled dataset of 45,507 species with missing trait records imputed using BHPMF (See Methods). The results of all PCAs show strong similarity, indicating robustness of the pattern obtained in the main analysis. Only PC1 and PC2 were identified as significant (see Methods) and therefore are reported here. All PCAs were performed on the correlation matrix of log₁₀-transformed traits.

Extended Data Table 2 | Description and illustrative examples of species at different positions at the margin of the global spectrum of plant form and function

Brief description and examples	
①	Tall, very large-seeded trees with large leaves of intermediate LMA and N_{mass} . Examples include the Neotropical <i>Bertholettia excelsa</i> (Brazil nut), <i>Gustavia superba</i> , <i>Pentaclethra macroloba</i> , and <i>Omphalea</i> spp.
②	Relatively large-seeded shrubs and trees of various heights, with small, sclerophyllous, highly conservative (high LMA, low N_{mass}) leaves. Examples include, among gymnosperms, monkey puzzle tree (<i>Araucaria araucana</i>), giant sequoia (<i>Sequoiadendron giganteum</i>) and junipers (<i>Juniperus</i> spp.). Among angiosperms, it includes members of the Proteaceae (e.g. the Australian <i>Hakea</i> and the South African <i>Leucadendron</i> genera) and Myrtaceae families (<i>Melaleuca uncinata</i> and <i>Eucalyptus dumosa</i>).
③	Sclerophyllous, high-LMA, low- N_{mass} species, of small stature, leaves, and seeds, varying from shrubs (such as <i>Fumana thymifolia</i> , heathers <i>Calluna vulgaris</i> and <i>Erica tetralix</i> , and chamise <i>Adenostoma fasciculatum</i>), to small forbs or sub-shrubs (such as <i>Diapensia lapponica</i> , <i>Draba</i> spp. and <i>Sedum</i> spp.), to graminoids (e.g. <i>Muhlenbergia ramulosa</i> and <i>Aristida purpurea</i>).
④	Submerged and semi-submerged aquatics (such as bladderwort <i>Utricularia vulgaris</i> , watermilfoil <i>Myriophyllum spicatum</i> , <i>Zannichellia palustris</i> , and <i>Ranunculus aquatilis</i>) and ephemeral, small-seeded and small- and acquisitive-leaved (low LMA, high N_{mass}) species of very short stature, with very low investment in vegetative structures other than leaves (such as thale cress <i>Arabidopsis thaliana</i> , annual bluegrass <i>Poa annua</i> , and <i>Nama dichotoma</i>).
⑤	Large-leaved, high- N_{mass} herbaceous plants with little carbon investment in support tissue. These are illustrated by robust aquatic species such as larger pondweeds (<i>Potamogeton</i> spp.) and sacred lotus (<i>Nelumbo nucifera</i>), species with nitrogen-rich secondary compounds (presumably anti-herbivore defences; e.g. devil's snare <i>Datura stramonium</i> , henbane <i>Hyoscyamus niger</i>), and some common crop and agricultural weed species such as <i>Beta vulgaris</i> , <i>Phaseolus vulgaris</i> , <i>Cannabis sativa</i> , and <i>Arctium minus</i> . Also includes Boreal 'mega-herb' <i>Angelica archangelica</i> .

Circled numbers in the first column refer to extreme poles of whole-plant specialization, whose approximate positions in the plane defined by PC1 and PC2 are indicated within circles in Fig. 2a.

An ID2-dependent mechanism for VHL inactivation in cancer

Sang Bae Lee¹, Veronique Frattini¹, Mukesh Bansal^{2,3}, Angelica M. Castano¹, Dan Sherman⁴, Keino Hutchinson⁴, Jeffrey N. Bruce⁵, Andrea Califano^{2,3}, Guangchao Liu¹, Timothy Cardozo⁴, Antonio Iavarone^{1,6,7} & Anna Lasorella^{1,7,8}

Mechanisms that maintain cancer stem cells are crucial to tumour progression. The ID2 protein supports cancer hallmarks including the cancer stem cell state. HIF α transcription factors, most notably HIF2 α (also known as EPAS1), are expressed in and required for maintenance of cancer stem cells (CSCs). However, the pathways that are engaged by ID2 or drive HIF2 α accumulation in CSCs have remained unclear. Here we report that DYRK1A and DYRK1B kinases phosphorylate ID2 on threonine 27 (Thr27). Hypoxia downregulates this phosphorylation via inactivation of DYRK1A and DYRK1B. The activity of these kinases is stimulated in normoxia by the oxygen-sensing prolyl hydroxylase PHD1 (also known as EGLN2). ID2 binds to the VHL ubiquitin ligase complex, displaces VHL-associated Cullin 2, and impairs HIF2 α ubiquitylation and degradation. Phosphorylation of Thr27 of ID2 by DYRK1 blocks ID2–VHL interaction and preserves HIF2 α ubiquitylation. In glioblastoma, ID2 positively modulates HIF2 α activity. Conversely, elevated expression of DYRK1 phosphorylates Thr27 of ID2, leading to HIF2 α destabilization, loss of glioma stemness, inhibition of tumour growth, and a more favourable outcome for patients with glioblastoma.

The HIF α (hypoxia inducible factor alpha) transcription factors are the key mediators of the hypoxia response¹. HIF α protein dysregulation in cancer can be triggered by mutation of the von-Hippel Lindau (VHL) gene, an event that hinders the negative control of HIF α protein stability through the ubiquitin ligase activity of VHL. This idea has been validated for HIF2 α , the HIF isoform preferentially upregulated in VHL-mutant tumours and which has recently been implicated as driver of cancer stem cells^{2–5}. However, signalling events that link the stem cell-intrinsic transcriptional machinery to pivotal mechanisms of HIF2 α regulation in cancer remain to be charted.

ID proteins (ID1 to ID4) are master regulators of stem cells that are hijacked during tumorigenesis and foster stem cell self-renewal and angiogenesis^{6,7}. Although the pro-tumorigenic role of ID proteins has been linked to the accumulation of mRNAs and proteins, it remains unclear whether other mechanisms exist that deregulate ID activity in cancer cells. Among ID proteins, ID2 is essential for tumour angiogenesis and glioma stemness and it is a component of the signature that predicts poor outcome in patients with high-grade glioma^{8–10}.

We show that ID2 activity is restrained by DYRK1 kinase-mediated phosphorylation on Thr27 and hypoxia reduces this event by inhibiting DYRK1 activity. In hypoxic brain tumour cells, active ID2 binds to and disrupts the VHL complex, thus preventing ubiquitin-mediated proteasomal degradation of HIF2 α . This is a previously unrecognized mechanism that fosters cancer stem cells and aggressiveness of human cancer.

Hypoxia regulates phosphorylation of ID2 by DYRK1

We used mass spectrometry to identify the phosphorylation sites of human ID2. Beside Ser5 (ref. 11), we found that ID2 is phosphorylated on Ser14 and Thr27 (Extended Data Fig. 1a–c). A sequencing analysis of the *ID2* gene in cancer revealed that the colorectal cancer cell line HRT-18 harbours and expresses a mutant ID2(T27A) protein (Extended Data Fig. 2a, b). Thr27 of ID2 is highly conserved throughout evolution (Extended Data Fig. 2c). The primary role of ID proteins is to preserve stem cell properties, a function widely

documented in neural stem cells (NSCs)^{12,13}. Therefore, to interrogate the significance of the ID2(T27A) mutation, we tested the self-renewing capacity of ID2-null NSCs reconstituted with wild-type ID2 (ID2(WT)) or ID2(T27A) (Fig. 1a). Introduction of ID2(T27A) in ID2-null NSCs increased neurosphere formation in serial passages by more than 50% when compared with ID2(WT) ($P = 0.00883$ – 0.000229 ; t ratio = 4.772 – 12.597) and caused a 2.4-fold increase in cell expansion rate (40.5 ± 1.7 versus 16.7 ± 0.831 ; $P < 0.0001$, Fig. 1b–d). From the analysis of 18 candidate kinases, the dual-specificity tyrosine-phosphorylation-regulated kinases 1A and 1B (DYRK1A and DYRK1B) emerged as the only enzymes able to phosphorylate Thr27 of ID2 (Fig. 1e, Extended Data Fig. 2d). The sequence surrounding the Thr27 residue in ID2 resembles the DYRK1 phosphorylation consensus motif RX(X)(S/T)P and is highly conserved in different species (Extended Data Fig. 2c)¹⁴. Antibodies against a phospho-T27-ID2 peptide confirmed that ID2 is phosphorylated by wild-type kinase but not the inactive DYRK1B(K140R) kinase (Fig. 1a, f–h)¹⁵. Endogenous and exogenous ID2 and ID2(T27A) co-precipitated endogenous DYRK1A and DYRK1B (Fig. 1i and Extended Data Fig. 2e). Treatment of glioma cells with harmine, a small-molecule inhibitor of DYRK1 (ref. 16), or combined short hairpin RNA (shRNA)-mediated silencing of DYRK1A and DYRK1B reduced Thr27 phosphorylation of ID2 (Fig. 3e and Extended Data Fig. 2f).

Next, we sought to identify the regulatory mechanisms controlling Thr27 phosphorylation of ID2. Exposure of human GBM-derived glioma stem cells (GSCs) to hypoxia or hypoxia-mimicking agent cobalt chloride (CoCl₂) caused loss of Thr27 phosphorylation (Fig. 2a and Extended Data Fig. 3a). Determination of the Thr27 phosphorylation stoichiometry of ID2 in the neuronal cell line SK-N-SN revealed that 21.08% of ID2 was phosphorylated on Thr27 in normoxia but the phosphorylation dropped to 2.28% in a hypoxic environment (Extended Data Fig. 3b, c). Mirroring the reduction of Thr27 phosphorylation of ID2, CoCl₂ reduced DYRK1 kinase activity (Fig. 2b and Extended Data Fig. 3d) and DYRK1 auto-phosphorylation, an event

¹Institute for Cancer Genetics, Columbia University Medical Center, New York 10032, USA. ²Department of Systems Biology, Columbia University Medical Center, New York 10032, USA. ³Center for Computational Biology and Bioinformatics, Columbia University Medical Center, New York 10032, USA. ⁴Department of Biochemistry and Molecular Pharmacology, NYU School of Medicine, New York 10014, USA. ⁵Department of Neurosurgery, Columbia University Medical Center, New York 10032, USA. ⁶Department of Neurology, Columbia University Medical Center, New York 10032, USA. ⁷Department of Pathology, Columbia University Medical Center, New York 10032, USA. ⁸Department of Pediatrics, Columbia University Medical Center, New York 10032, USA.

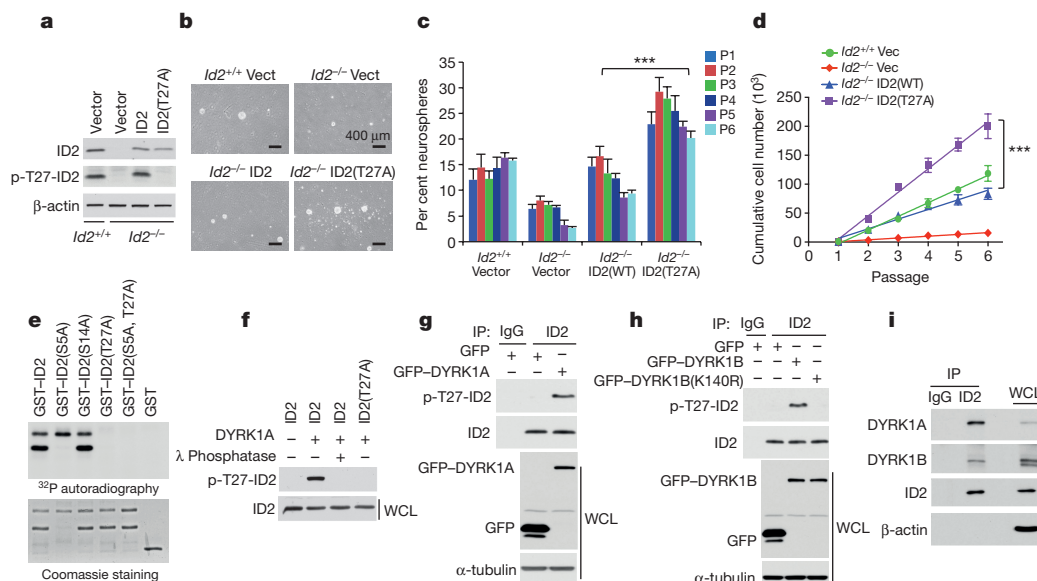


Figure 1 | Inhibition of DYRK1-mediated phosphorylation of ID2 at Thr27 promotes NSC properties. **a**, Reconstitution of NSCs $ID2^{-/-}$ with ID2(WT), ID2(T27A) or the empty vector. **b**, Microphotographs of representative cultures from the neurosphere-forming assay. **c**, Percent of neurospheres generated in serial clonal assays (means of 3 biological replicates \pm s.d.; *** $P=0.00883-0.000229$ for $ID2^{-/-}$ ID2(T27A) compared with $ID2^{-/-}$ ID2(WT)). **d**, Cumulative cell number of cultures as in **c** (means of 3 biological replicates \pm s.d. *** $P<0.0001$ for $ID2^{-/-}$

ID2(T27A) compared with $ID2^{-/-}$ ID2(WT)). **e**, *In vitro* kinase assay shows phosphorylation of GST-ID2 protein by recombinant DYRK1B. **f**, Phosphorylation of ID2 but not ID2(T27A) by DYRK1A in IMR32 cells. **g**, Phosphorylation of endogenous ID2 by DYRK1A in U87 cells. **h**, Phosphorylation of endogenous ID2 by DYRK1B but not the kinase inactive DYRK1B(K140R) in U87 cells. **i**, Binding between endogenous DYRK1A or DYRK1B and ID2. WCL, whole cellular lysate.

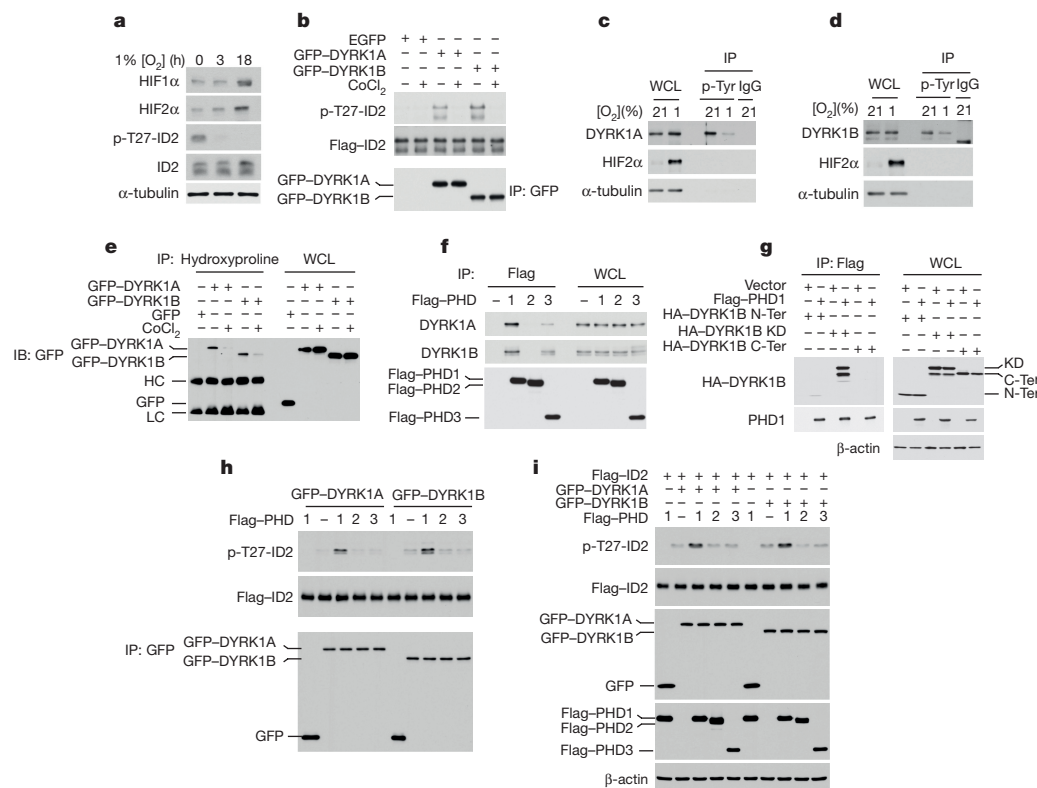


Figure 2 | DYRK1 kinases and ID2 Thr27 phosphorylation are inhibited by hypoxia and enhanced by PHD1. **a**, Hypoxia inhibits phosphorylation of ID2 Thr27 in GSC line 1123. **b**, $CoCl_2$ inhibits DYRK1 kinase. GFP-DYRK1 immunoprecipitates from U87 cells untreated or treated with $CoCl_2$ and recombinant Flag-ID2 were used in kinase assay *in vitro*. **c**, **d**, Hypoxia reduces tyrosine phosphorylation of DYRK1A (**c**) and DYRK1B (**d**) as evaluated by anti-p-Tyr immunoprecipitation in GSC line 1123 cells. **e**, $CoCl_2$ inhibits proline hydroxylation of DYRK1A and

DYRK1B as shown by anti-hydroxyproline immunoprecipitation in U87 glioma cells. (HC, IgG heavy chain; LC, IgG light chain). **f**, Endogenous DYRK1A and DYRK1B interact with Flag-PHD1 in U87 cells. **g**, The kinase domain (KD) but not the N- or the C-terminal domains of DYRK1B interacts with PHD1 in co-immunoprecipitation assay. **h**, Expression of PHD1 enhances cellular DYRK1 kinase activity in an *in vitro* phosphorylation assay using recombinant ID2. **i**, Expression of PHD1 enhances DYRK1 kinase activity towards ID2 Thr27 *in vivo*.

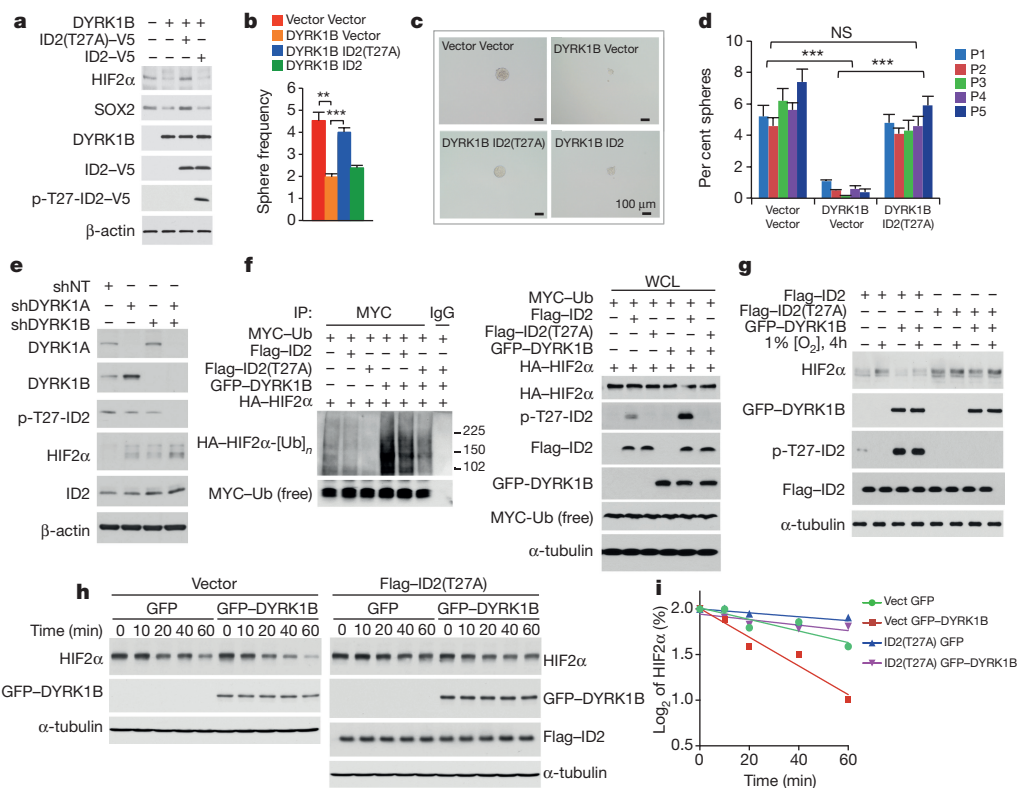


Figure 3 | The DYRK1-ID2-Thr27 pathway controls GSCs and HIF2 α . **a**, Phosphorylation of ID2 but not ID2(T27A) by GFP-DYRK1B downregulates HIF2 α and SOX2 in GSC line 31. **b**, *In vitro* LDA of parallel cultures shows that decreased frequency of gliomaspheres by DYRK1B is rescued by ID2(T27A). Data are means of 3 biological replicates \pm s.d.; ** P = 0.0031 (vector versus DYRK1B); *** P = 0.00022 (DYRK1B versus DYRK1B plus ID2(T27A)). **c**, Microphotographs of representative cultures in **b**. **d**, Serial clonal experiments of cells in **b**. Data are means of 3 biological replicates \pm s.d. of percent gliomaspheres; *** P = 0.00059–0.00007 for vector versus DYRK1B plus vector; *** P = 0.0089–0.0008 for ID2(T27A) plus DYRK1B versus DYRK1B plus

vector; NS: P = 0.061–0.249 for ID2(T27A) plus DYRK1B versus vector. **e**, Silencing of DYRK1 downregulates phospho-Thr27 of ID2 and increases HIF2 α in U87 cells. **f**, Ubiquitylation of HIF2 α is enhanced by DYRK1B and reduced by ID2(T27A) as evaluated by *in vivo* ubiquitylation (left panels, MYC-Ub immunoprecipitation/HA-HIF2 α western blot; right panels, whole cellular lysates, WCL). **g**, ID2(T27A) elevates HIF2 α and opposes DYRK1B-mediated reduction of HIF2 α during hypoxia. **h**, ID2(T27A) reverts DYRK1B-mediated decrease of HIF2 α half-life during recovery from exposure to CoCl₂. U87 were exposed to CoCl₂ for 3 h followed by CoCl₂ washout. **i**, Quantification of HIF2 α protein from the experiment in **h**.

required for the activity of DYRK1 kinase¹⁷ (Extended Data Fig. 3e–g). Similarly, exposure of GSCs to low oxygen decreased DYRK1A and DYRK1B tyrosine auto-phosphorylation (Fig. 2c, d).

Prolyl hydroxylases PHD1, PHD2 and PHD3 operate as direct sensors of cellular oxygen concentration^{18,19}. Immunoprecipitation using an antibody that recognizes hydroxyprolines indicated that DYRK1A and DYRK1B carry hydroxylated prolines, and CoCl₂ abrogated DYRK1 prolyl hydroxylation (Fig. 2e). DYRK1A and DYRK1B interacted *in vivo* with PHD1 (Fig. 2f) and the expression of PHD1 enhanced prolyl hydroxylation of both kinases (Extended Data Fig. 3h). In particular, DYRK1B interacted with PHD1 through the kinase domain (Fig. 2g). The activity of DYRK1A and DYRK1B towards Thr27 of ID2 was potentiated by PHD1 *in vitro* and *in vivo* (Fig. 2h, i). Thus, oxygen deprivation induces a constitutively active ID2 by inactivating DYRK1 kinases, which are positively regulated substrates of PHD1.

Phosphorylation of ID2 by DYRK1 destabilizes HIF2 α

We used human GSCs to interrogate the effects of DYRK1 and ID2(T27A) on HIF2 α and glioma stemness. Lentiviral transduction of the DYRK1-resistant ID2(T27A) mutant in GSC line number 48 resulted in elevation of HIF2 α and enhanced tumour sphere forming capacity as measured by limiting dilution assay (LDA, Extended Data Fig. 4a–d). ID2(T27A)-induced accumulation of the HIF2 α protein was independent of transcription (Extended Data Fig. 4e). When detectable, HIF1 α levels mirrored those of HIF2 α but with more limited changes (Extended Data Fig. 4f).

Expression of DYRK1 in GSC line number 34 and GSC line number 31 reduced HIF2 α , the HIF2 α target TGF α and the glioma stem cell marker SOX2 (Fig. 3a and Extended Data Fig. 4g, h). Also in this case HIF2 α mRNA was unchanged (Extended Data Fig. 4h, j). LDA and serial clonal experiments showed that the DYRK1-induced decrease of HIF2 α attenuated glioma stemness (Fig. 3b–d and Extended Data Fig. 4k). However, accumulation of HIF2 α , expression of SOX2 and the frequency of GSCs were restored by co-expression of DYRK1 and ID2(T27A) but not ID2(WT) (Fig. 3a–d and Extended Data Fig. 4k). DYRK1-mediated inhibition of glioma sphere formation was overridden by co-expression of non-degradable HIF2 α (HIF2 α -TM, TM, triple mutant; Extended Data Fig. 4l)²⁰. Furthermore, silencing of DYRK1A or DYRK1B upregulated HIF2 α and reduced phosphorylation of Thr27 of ID2, with maximal effects after co-silencing of both kinases (Fig. 3e). Next, we investigated the effects of DYRK1 and ID2(T27A) on ubiquitylation and the stability of HIF α . DYRK1-mediated phosphorylation of Thr27 triggered HIF α ubiquitylation and expression of ID2(T27A) reversed DYRK1 effect (Fig. 3f and Extended Data Fig. 5a). Similarly, expression of DYRK1B prevented accumulation of HIF2 α under hypoxia and co-expression of ID2(T27A) abrogated this response (Fig. 3g). DYRK1 accelerated the decay of HIF2 α during recovery from exposure to CoCl₂ and reduced HIF2 α half-life and ID2(T27A) counteracted these effects (Fig. 3h, i and Extended Data Fig. 5b, c).

ID2 binds and disrupts the VCB-Cul2 complex

Mass spectrometry analysis of ID2 immunoaffinity complexes revealed that Elongin C, a component of the VCB-Cul2 ubiquitin ligase

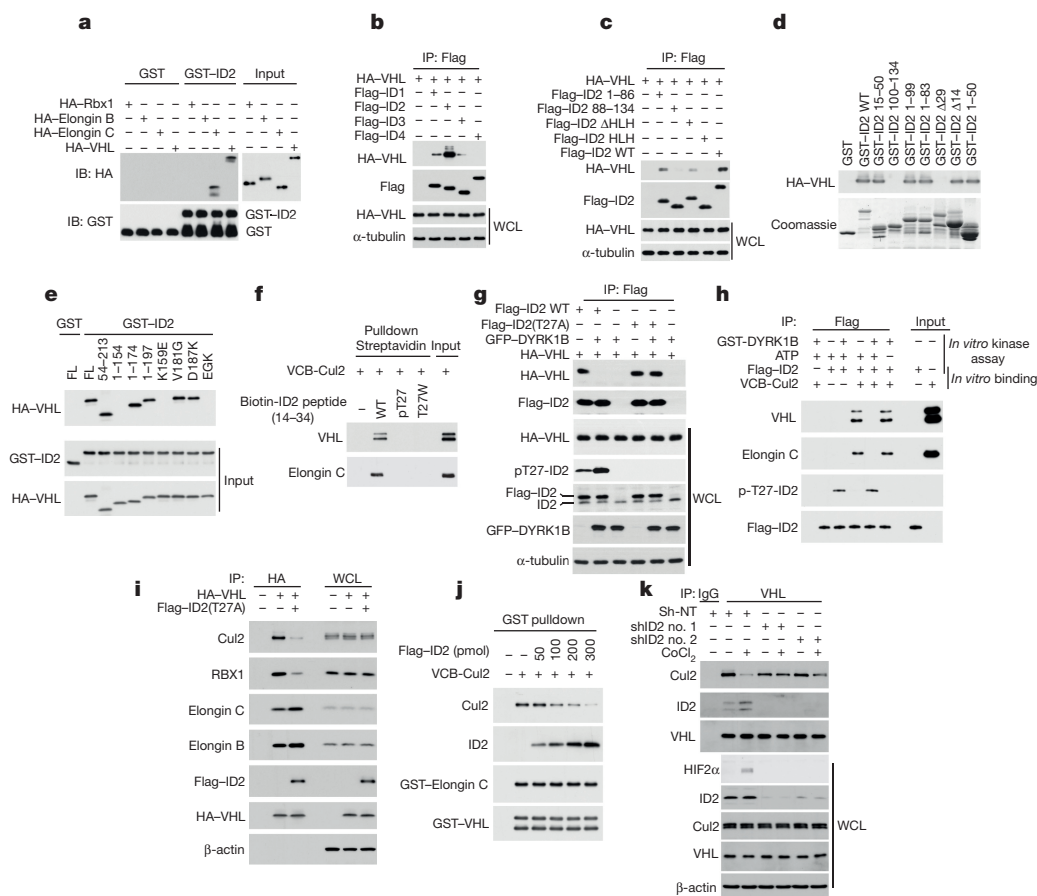


Figure 4 | Unphosphorylated ID2 binds to and disrupts the VCB-Cul2 complex. **a**, Recombinant ID2 interacts with Elongin C and VHL in a GST pull-down assay (input: 10%). **b**, Co-immunoprecipitation shows that VHL preferentially interacts with ID2. **c**, The N terminus of ID2 is required for the interaction with VHL as determined by immunoprecipitation and western blot (IP-WB) analysis. **d**, The N terminus of GST-ID2 is required for the interaction with *in vitro* translated HA-VHL. **e**, Amino acids 154–174 of *in vitro* translated HA-VHL mediate the interaction with GST-ID2 (input: 10%). **f**, Phosphorylation of ID2 Thr27 or the ID2(T27W) mutation disrupt the ID2-VHL interaction as analysed by *in vitro* streptavidin pull down of biotinylated ID2 peptides in the

presence of recombinant VCB-Cul2. **g**, Co-immunoprecipitation shows that DYRK1B-mediated phosphorylation of ID2 disrupts ID2 interaction with VHL *in vivo*. **h**, *In vitro* phosphorylation of recombinant ID2 by purified DYRK1B blocks ID2 interaction with VHL and Elongin C in the reconstituted VCB-Cul2 complex. **i**, ID2(T27A) displaces Cul2 from VCB complex in co-immunoprecipitation assays in U87 cells. **j**, Progressive dissociation of Cul2 from recombinant VCB complex by increasing concentration of purified Flag-ID2. **k**, Silencing of ID2 in U87 reverts CoCl₂-mediated dissociation of Cul2 from VHL as evaluated by Co-IP-WB. WCL, whole cellular lysates.

complex that includes VHL, Elongin C, Elongin B, Cullin 2, and RBX1 is an ID2-associated protein (Supplementary Tables 1 and 2). The direct interaction of ID2 with Elongin C and VHL was confirmed *in vitro* and *in vivo* (Fig. 4a and Extended Data Fig. 5d, e). ID2 was unable to bind to HIF α proteins (Extended Data Fig. 5d). VHL and Elongin C interacted strongly with ID2, weakly with ID1 and ID3, and did not bind to ID4 (Fig. 4b and Extended Data Fig. 5f, g). The interaction between ID2 and VHL was mediated by the amino-terminal region of ID2 that includes Thr27 and did not require the HLH domain (amino acids 35–76, Fig. 4c). A more detailed mapping of the regions involved in the VHL-ID2 interaction revealed that amino acids 15–35 of ID2 and the SOCS box of VHL (amino acids 154–174) were required for the binding (Fig. 4d, e). Mutation of Lys159 (K159E), which provides the VHL contact surface for binding to Cul2 (refs 21, 22), impaired the interaction with ID2 (Fig. 4e). Consistent with ID2 deletion mapping analysis, an ID2 peptide composed of amino acids 14–34 bound to both VHL and components of the VCB complex pre-assembled in insect cells. However, addition of a phosphate to Thr27 or mutation of Thr27 to a bulky hydrophobic amino acid (T27W) prevented the binding to both VHL and the VCB complex (Fig. 4f and Extended Data Fig. 5h). These findings were corroborated by computational molecular docking whereby a N-terminally-derived ID2 peptide (amino acids 15–31) docked preferentially to a groove on the molecular surface of

VHL:Elongin C²² with the N-terminal half of its interaction surface contacting the SOCS box of VHL that binds Cul2 (primarily Lys159) and the C-terminal half (including Thr27) fitting snugly into a hydrophobic pocket mostly contributed by the Elongin C surface (Extended Data Fig. 6a, c, d). Mutating Thr27 to phospho-Thr27 and re-docking resulted in unfavourable energy and displacement of the peptide from this location on the complex (Extended Data Fig. 6b). DYRK1 disrupted the interaction between VHL and ID2(WT) but did not affect the binding of VHL to ID2(T27A) (Fig. 4g). The phosphomimic ID2(T27E) mutant failed to bind VHL and did not promote accumulation of HIF2 α (Extended Data Fig. 7a, b). Next, we developed an *in vitro* assay using purified proteins that included bacterially expressed Flag-ID2, enzymatically active recombinant DYRK1B and baculovirus-expressed VCB-Cul2 complex or purified VHL. In this system, ID2 bound to VCB-Cul2 complex and VHL in the absence of active DYRK1B, but the interaction was disrupted by DYRK1-mediated phosphorylation of Thr27 (Fig. 4h and Extended Data Fig. 7c).

In the VCB-Cul2 complex, the Cul2 subunit provides the scaffold module for the interaction with the ubiquitin-conjugating enzyme (E2)^{23,24}. As we expressed ID2(T27A), ID2 was loaded onto VCB, and the Cul2/RBX1 module dissociated from the complex in the absence of changes in the total cellular levels of Cul2 and RBX1 (Fig. 4i). We also found that challenging the pre-assembled VCB complex with increasing

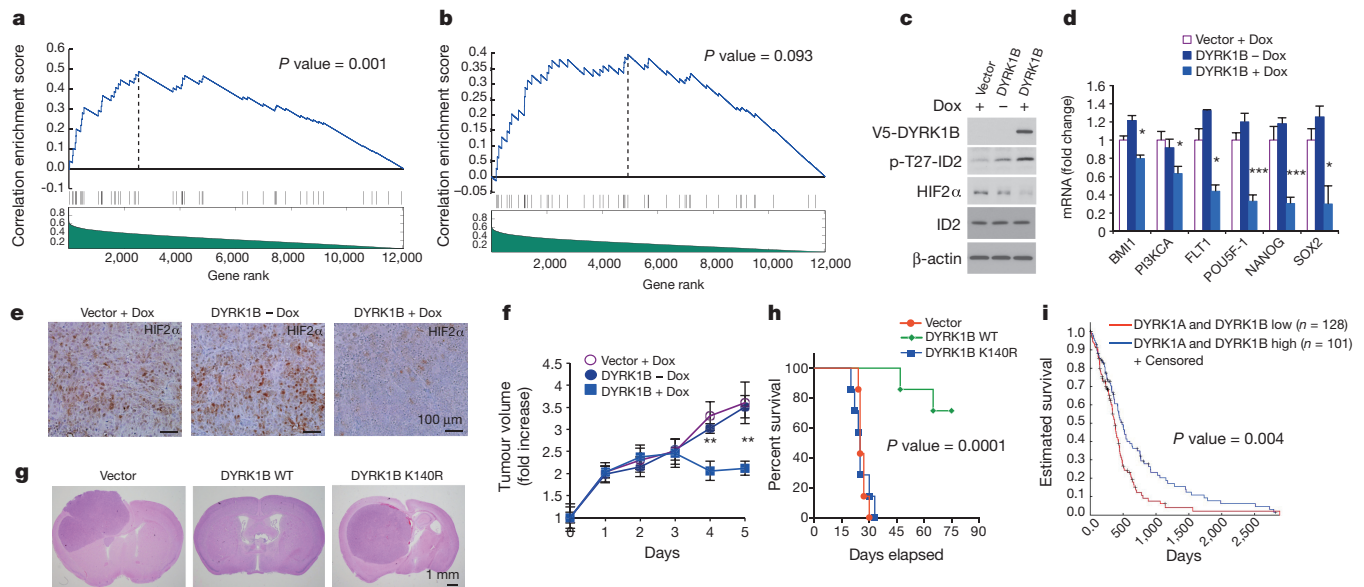


Figure 5 | DYRK1 kinases inhibit human glioma growth by repressing an ID2–HIF2 α network. **a**, Significant and positive targets of HIF2 α correlate with HIF2 α in GBM with high ID2 activity compared to a set of random genes by GSEA; **b**, Correlations are not significant in GBM with low ID2 activity **c**, Inducible expression of DYRK1B in U87 causes ID2 Thr27 phosphorylation and downregulation of HIF2 α . **d**, qRT-PCR from cells treated as in **c**; $n=9$ (3 biological replicates performed in triplicates) \pm s.d. *BMI1*: $*P=0.0470$; *PI3KCA*: $*P=0.0279$; *FLT1*: $*P=0.0246$; *POU5F-1*: $***P=0.000796344$; *NANOG*: $***P=0.000737396$; *SOX2*: $*P=0.028884239$ (DYRK1B minus Dox versus DYRK1B plus Dox). **e**, Inducible expression of DYRK1B

amounts of recombinant Flag–ID2 resulted in progressive dissociation of Cul2 (Fig. 4j). Expression of ID2(T27A) triggered a comparable block of Elongin C–Cul2 association whereas it did not affect the assembly of a Cul5-based complex containing SOCS2, a SOCS protein that cannot bind to ID2(T27A) (Extended Data Fig. 7d, e). Finally, in glioma cells in which ID2 and VHL are present at a molar ratio of 5.7:1 (Extended Data Fig. 7f), hypoxia signalling promoted the association between endogenous VHL and ID2 while dissociating Cul2 (Fig. 4k and Extended Data Fig. 7g). Silencing of ID2 rescued the dissociation of Cul2 from VCB complex and prevented HIF2 α elevation (Fig. 4k). Together, these findings indicate that ID2 activation stabilizes HIF2 α by disabling VCB ubiquitin ligase via dissociation of Cul2.

A DYRK1–ID2 pathway controls HIF2 α in glioma

To determine whether activation of ID2 enhances HIF2 α transcriptional activity in an unbiased fashion, we used CINDy, an algorithm for high-fidelity reconstruction of post-translational causal dependencies to interrogate whether ID2 can affect the activity of HIF2 α on its targets in the context of GBM²⁵. When applied to a collection of 548 TCGA-derived GBM samples, ID2 activity emerged as the modulator of the transcriptional connection between HIF2 α and its activated target genes (Fig. 5a, b). The activity of ID2 is estimated by the VIPER algorithm, a computational tool designed to infer protein activity from gene expression data²⁶. When GBM samples were divided into two groups based on ID2 activity, samples with higher ID2 activity showed significantly stronger correlation between HIF2 α and its targets than a set of random genes ($P=0.001$, Fig. 5a). This positive correlation was absent in the cohort of GBM with low ID2 activity ($P=0.093$, Fig. 5b). Consistent with these observations, we detected a marked reduction of HIF2 α protein following acute deletion of the *Id1* and *Id2* genes in a mouse model of malignant glioma (Extended Data Fig. 8a)⁹. We interrogated the effect of DYRK1 expression in mouse models of human glioma. Tetracycline-induced expression of DYRK1B at levels comparable to normal brain (Extended Data Fig. 8b), downregulated HIF2 α

downregulates HIF2 α in subcutaneous xenografts of U87 cells as indicated by immunostaining. **f**, Inducible expression of DYRK1B causes tumour growth inhibition in mice treated as in **e**. Doxycycline (Dox) treatment starts at day 0 ($n=7$ mice per group; $**P=0.0040$ and 0.0069 , DYRK1B – Dox versus DYRK1B + Dox at day 4 and day 5, respectively). **g**, Expression of DYRK1B(WT) but not DYRK1B(K140R) inhibits orthotopic growth of U87 (haematoxylin & eosin staining of brain cross-sections). Mice injected with U87-vector or DYRK1B(K140R) were euthanized on day 25. Mice injected with U87-DYRK1B were euthanized on day 70. **h**, Kaplan–Meier analysis of mice in **g** ($n=7$ animals per group). **i**, Expression of DYRK1A and DYRK1B predicts survival in GBM patients.

in glioma cells *in vitro* and in sub-cutaneous xenografts and reduced the expression of the HIF2 α targets that promote stem cell functions (Fig. 5c–e). Expression of DYRK1B also inhibited tumour cell proliferation *in vivo*, resulting in tumour reduction (Fig. 5f and Extended Data Fig. 8c, d). Next, we evaluated the anti-tumour effects of DYRK1B(WT) or the kinase inactive K140R mutant in an orthotopic model of glioma (Extended Data Fig. 8e). Animals bearing glioma cells that expressed DYRK1B(WT) manifested significantly increased survival and tumour latency relative to mice bearing DYRK1B(K140R) or vector transduced cells (Fig. 5g, h). Two out of seven mice in the DYRK1B(WT) group developed tumours that failed to express exogenous DYRK1B (Extended Data Fig. 8f). This result suggests that active DYRK1 kinase is incompatible with tumour growth in this glioma model. Finally, higher DYRK1A and DYRK1B predicted a more favourable clinical outcome for GBM patients, thus supporting the clinical significance of DYRK1 activity in glioma (Fig. 5i and Extended Data Fig. 9a).

Discussion

Here we report a novel mechanism of functional inactivation of the VHL ubiquitin ligase that is independent of genetic mutations of the *VHL* gene. This mechanism is centred on the ability of active ID2 to disrupt the VCB–Cul2 complex, leading to HIF2 α stabilization. We also unravel a hypoxia-directed cascade of events that by overriding the restraining effect of DYRK1-mediated phosphorylation of ID2 on Thr27, culminates with ID2 activation and HIF2 α stabilization. The transcriptional activation of the *ID2* gene, a HIF α target²⁷, by HIF2 α generates a feed-forward ID2–HIF2 α loop that further supports cancer stem cells and glioma aggressiveness (Extended Data Fig. 10). By showing that PHD1-mediated prolyl hydroxylation enhances the enzymatic activity of DYRK1 kinases towards ID2, our findings provide a clue to the mechanism of DYRK1 inhibition in hypoxia. Thus, inhibition of DYRK1 kinases is an oxygen-sensing signal that disables VCB–Cul2.

The gene coding for *DYRK1A* is gained in Down syndrome, a disease characterized by impaired neural proliferation during

development^{28–30}, reduced self-renewal and premature withdrawal from the cell cycle³¹. These findings are consistent with the notion that activation of DYRK1A and DYRK1B inhibits proliferation and activates a cellular quiescence program^{32–37}. It is plausible that DYRK1-activating signals are negative regulators of self-renewal of both normal and cancer stem cells. By phosphorylating ID2 on Thr27, DYRK1 weakens the core of the stemness machinery centred on the ID2–HIF2 α pathway. We suggest that together with the control of ID2 activity, regulation of sonic hedgehog/Gli signalling and CREB-mediated transcription by DYRK kinases cooperate to restrain stem cell functions and tumorigenesis in the nervous system^{38–40}. The glioma inhibitory effect of DYRK1 in the mouse, and the predictive value of DYRK1 expression on GBM patient survival support the tumour suppressor function of DYRK1 kinases *in vivo*. Whereas there is agreement on the negative role of DYRK1 towards stemness and cell proliferation^{31,35}, the broad spectrum of DYRK kinase substrates may account for the pro-tumorigenic functions of DYRK1 reported by other studies⁴¹. It has been proposed that loss of DYRK1 and accumulation of ID2 and HIF2 α drive tumour progression of other cancer types beside GBM. The mechanism reported here might operate widely during the transition of cancer cells towards the more aggressive stem-like state that drives maintenance and progression of solid tumours.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 April; accepted 23 November 2015.

Published online 6 January 2016.

- Keith, B., Johnson, R. S. & Simon, M. C. HIF1 α and HIF2 α : sibling rivalry in hypoxic tumour growth and progression. *Nature Rev. Cancer* **12**, 9–22 (2012).
- Gordan, J. D., Bertout, J. A., Hu, C. J., Diehl, J. A. & Simon, M. C. HIF-2 α promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity. *Cancer Cell* **11**, 335–347 (2007).
- Kondo, K., Kim, W. Y., Lechpammer, M. & Kaelin, W. G., Jr. Inhibition of HIF2 α is sufficient to suppress pVHL-defective tumor growth. *PLoS Biol.* **1**, E83 (2003).
- Li, Z. *et al.* Hypoxia-inducible factors regulate tumorigenic capacity of glioma stem cells. *Cancer Cell* **15**, 501–513 (2009).
- Pietras, A. *et al.* HIF-2 α maintains an undifferentiated state in neural crest-like human neuroblastoma tumor-initiating cells. *Proc. Natl Acad. Sci. USA* **106**, 16805–16810 (2009).
- Lasorella, A., Benezra, R. & Iavarone, A. The ID proteins: master regulators of cancer stem cells and tumour aggressiveness. *Nature Rev. Cancer* **14**, 77–91 (2014).
- Perk, J., Iavarone, A. & Benezra, R. Id family of helix-loop-helix proteins in cancer. *Nature Rev. Cancer* **5**, 603–614 (2005).
- Lasorella, A., Rothschild, G., Yokota, Y., Russell, R. G. & Iavarone, A. Id2 mediates tumor initiation, proliferation, and angiogenesis in Rb mutant mice. *Mol. Cell. Biol.* **25**, 3563–3574 (2005).
- Niola, F. *et al.* Mesenchymal high-grade glioma is maintained by the ID–RAP1 axis. *J. Clin. Invest.* **123**, 405–417 (2013).
- Vandeputte, D. A. *et al.* Expression and distribution of id helix-loop-helix proteins in human astrocytic tumors. *Glia* **38**, 329–338 (2002).
- Hara, E., Hall, M. & Peters, G. Cdk2-dependent phosphorylation of Id2 modulates activity of E2A-related transcription factors. *EMBO J.* **16**, 332–342 (1997).
- Nam, H. S. & Benezra, R. High levels of Id1 expression define B1 type adult neural stem cells. *Cell Stem Cell* **5**, 515–526 (2009).
- Niola, F. *et al.* Id proteins synchronize stemness and anchorage to the niche of neural stem cells. *Nature Cell Biol.* **14**, 477–487 (2012).
- Himpel, S. *et al.* Specificity determinants of substrate recognition by the protein kinase DYRK1A. *J. Biol. Chem.* **275**, 2431–2438 (2000).
- Lee, K., Deng, X. & Friedman, E. Mirk protein kinase is a mitogen-activated protein kinase substrate that mediates survival of colon cancer cells. *Cancer Res.* **60**, 3631–3637 (2000).
- Göckler, N. *et al.* Harmine specifically inhibits protein kinase DYRK1A and interferes with neurite formation. *FEBS J.* **276**, 6324–6337 (2009).
- Himpel, S. *et al.* Identification of the autophosphorylation sites and characterization of their effects in the protein kinase DYRK1A. *Biochem. J.* **359**, 497–505 (2001).
- Kaelin, W. G. Jr & Ratcliffe, P. J. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Mol. Cell* **30**, 393–402 (2008).
- Semenza, G. L. HIF-1, O(2), and the 3 PHDs: how animal cells signal hypoxia to the nucleus. *Cell* **107**, 1–3 (2001).
- Warnecke, C. *et al.* Differentiating the functional role of hypoxia-inducible factor (HIF)-1 α and HIF-2 α (EPAS-1) by the use of RNA interference: erythropoietin is a HIF-2 α target gene in Hep3B and Kelly cells. *FASEB J.* **18**, 1462–1464 (2004).
- Kershaw, N. J. & Babon, J. J. VHL: cullin-g the hypoxic response. *Structure* **23**, 435–436 (2015).
- Nguyen, H. C., Yang, H., Fribourgh, J. L., Wolfe, L. S. & Xiong, Y. Insights into Cullin-RING E3 ubiquitin ligase recruitment: structure of the VHL–EloBC–Cul2 complex. *Structure* **23**, 441–449 (2015).
- Kamura, T. *et al.* Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase. *Science* **284**, 657–661 (1999).
- Ohta, T., Michel, J. J., Schottelius, A. J. & Xiong, Y. ROC1, a homolog of APC11, represents a family of cullin partners with an associated ubiquitin ligase activity. *Mol. Cell* **3**, 535–541 (1999).
- Giorgi, F. M. *et al.* Inferring protein modulation from gene expression data using conditional mutual information. *PLoS ONE* **9**, e109569 (2014).
- Alvarez, M. J., Giorgi, F. M. & Califano, A. Using viper, a package for virtual inference of protein-activity by enriched regulon analysis. *Bioconductor*, 1–14 (2014).
- Löfstedt, T. *et al.* Induction of ID2 expression by hypoxia-inducible factor-1: a role in dedifferentiation of hypoxic neuroblastoma cells. *J. Biol. Chem.* **279**, 39223–39231 (2004).
- Chakrabarti, L., Galdzicki, Z. & Haydar, T. F. Defects in embryonic neurogenesis and initial synapse formation in the forebrain of the Ts65Dn mouse model of Down syndrome. *J. Neurosci.* **27**, 11483–11495 (2007).
- Contestabile, A. *et al.* Cell cycle alteration and decreased cell proliferation in the hippocampal dentate gyrus and in the neocortical germinal matrix of fetuses with Down syndrome and in Ts65Dn mice. *Hippocampus* **17**, 665–678 (2007).
- Guimeá, J. *et al.* A human homologue of *Drosophila* minibrain (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical region. *Hum. Mol. Genet.* **5**, 1305–1310 (1996).
- Hibaoui, Y. *et al.* Modelling and rescuing neurodevelopmental defect of Down syndrome using induced pluripotent stem cells from monozygotic twins discordant for trisomy 21. *EMBO Mol. Med.* **6**, 259–277 (2014).
- Aranda, S., Laguna, A. & de la Luna, S. DYRK family of protein kinases: evolutionary relationships, biochemical properties, and functional roles. *FASEB J.* **25**, 449–462 (2011).
- Becker, W. Emerging role of DYRK family protein kinases as regulators of protein stability in cell cycle control. *Cell Cycle* **11**, 3389–3394 (2012).
- Hämmerle, B. *et al.* Transient expression of Mnb/Dyrk1a couples cell cycle exit and differentiation of neuronal precursors by inducing p27KIP1 expression and suppressing NOTCH signaling. *Development* **138**, 2543–2554 (2011).
- Litovchick, L., Florens, L. A., Swanson, S. K., Washburn, M. P. & DeCaprio, J. A. DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. *Genes Dev.* **25**, 801–813 (2011).
- Park, J. *et al.* Dyrk1A phosphorylates p53 and inhibits proliferation of embryonic neuronal cells. *J. Biol. Chem.* **285**, 31895–31906 (2010).
- Yabut, O., Domogauer, J. & D'Arcangelo, G. Dyrk1A overexpression inhibits proliferation and induces premature neuronal differentiation of neural progenitor cells. *J. Neurosci.* **30**, 4004–4014 (2010).
- Lauth, M. *et al.* DYRK1B-dependent autocrine-to-paracrine shift of Hedgehog signaling by mutant RAS. *Nature Struct. Mol. Biol.* **17**, 718–725 (2010).
- Mao, J. *et al.* Regulation of Gli1 transcriptional activity in the nucleus by Dyrk1. *J. Biol. Chem.* **277**, 35156–35161 (2002).
- Yang, E. J., Ahn, Y. S. & Chung, K. C. Protein kinase Dyrk1 activates cAMP response element-binding protein during neuronal differentiation in hippocampal progenitor cells. *J. Biol. Chem.* **276**, 39819–39824 (2001).
- Pozo, N. *et al.* Inhibition of DYRK1A destabilizes EGFR and reduces EGFR-dependent glioblastoma growth. *J. Clin. Invest.* **123**, 2475–2487 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank N. Sherman for phospho-ID2 and ID2-associated proteins analysis by mass spectrometry, C. Warnecke for the plasmid expressing HIF2 α -TM, Z. Ronai for Flag-PHD1, Flag-PHD2, and Flag-PHD3, plasmids, A. Flores-Morales for the plasmid expressing Flag-SOCS2, M. Pagano for cDNAs for RBX1, Elongin B, Elongin C, and K. H. Kim for pcDNA-VHL. We thank D. D'Arca for preparation of VHL and HIF α constructs. This work was supported by National Institute of Health grants to A. L. (R01CA101644 and R01CA131126), and A. I. (R01CA178546 and R01NS061776 and a grant from The Chemotherapy Foundation). V. F. is supported by a fellowship from the American Brain Tumor Association (ABTA). S. B. L. was supported by NRF-2013R1A6A3A03063888 fellowship.

Author Contributions A.I. and A.L. conceived the project, coordinated the study, oversaw the results and wrote the manuscript. S.B.L. designed and performed most biochemical and cell biology experiments and helped with writing the manuscript; V.F. generated lentiviral vectors, performed *in vitro* GSC infections, and RT-PCR and binding studies; A.M.C. generated and analysed glioma xenografts and assisted in GSC experiments; G.L. conducted the *in vitro* screening assay that identified DYRK1 as the ID2 Thr27 kinase. M.B. and A.C. conducted gene expression and bioinformatics analyses. J.N.B. provided excess tissue from human GBM for GSC isolation. D.S., K.H. and T.C. generated the computational molecular docking model. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.L. (al2179@columbia.edu) or A.I. (ai2102@columbia.edu).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment.

Plasmids, cloning and lentivirus production. A constitutively stabilized mutant of HIF2 α (HIF2 α -TM) was obtained from Christina Warnecke²⁰. The HIF2 α -TM (triple mutant) construct harbours the following mutations in the prolyl and asparagyl hydroxylation sites: P405A, P530G and N851A. Polypeptide fragments of DYRK1B were cloned into pcDNA3-HA and include DYRK1B N terminus, N-Ter (amino acids 1–110), DYRK1B kinase domain, KD (amino acids 111–431), and DYRK1B C terminus, C-Ter (amino acids 432–629). cDNAs for RBX1, Elongin B and Elongin C were kindly provided from Michele Pagano (New York University) and cloned into the pcDNA vector by PCR. HA-tagged HIF1 α and HIF2 α were obtained from Addgene. GFP-tagged DYRK1A and DYRK1B were cloned into pcDNA vector. pcDNA-HA-VHL was provided by Kook Hwan Kim (Sungkyunkwan University School of Medicine, Korea). Site-directed mutagenesis was performed using QuickChange or QuickChange Multi Site-Directed mutagenesis kit (Agilent) and resulting plasmids were sequence verified. Lentivirus was generated by co-transfection of the lentiviral vectors with pCMV- Δ R8.1 and pMD2.G plasmids into HEK293T cells as previously described^{9,42}. ShRNA sequences are: ID2-1: GCCTACTGAATGCTGTATACCTGAGTATACACAGCATTCAGTAGGC; ID2-2: CCCACTATTGTGTCAGCCTGCATCTCGAGATGCAGGCTGACA ATAGTGGG; DYRK1A: CAGGTTGTAAAGGCATATGATCTCGAGATC ATATGCCTTTACAACCTG; DYRK1B: GACCTACAAGCACATCAATGA CTCGAGTCATTGATGTGCTTGTAGGTC.

Cell culture and hypoxia induction. IMR-32 (ATCC CCL-127), SK-N-SH (ATCC HTB-11), U87 (ATCC HTB-14), NCI-H1299 (ATCC CRL-5803), HRT18 (ATCC CCL-244), and HEK293T (ATCC CRL-11268) cell lines were acquired through American Type Culture Collection. U251 (Sigma, catalogue number 09063001) cell line was obtained through Sigma. Cell lines were cultured in DMEM supplemented with 10% fetal bovine serum (FBS, Sigma). Cells were routinely tested for mycoplasma contamination using Mycoplasma Plus PCR Primer Set (Agilent, Santa Clara, CA) and were found to be negative. Cells were transfected with Lipofectamine 2000 (Invitrogen) or calcium phosphate. Mouse NSCs were grown in Neurocult medium (StemCell Technologies) containing 1 \times proliferation supplements (StemCell Technologies), and recombinant FGF-2 and EGF (20 ng ml⁻¹ each; Peprotech). GBM-derived glioma stem cells were obtained by de-identified brain tumour specimens from excess material collected for clinical purposes at New York Presbyterian-Columbia University Medical Center. Donors (patients diagnosed with glioblastoma) were anonymous. Progressive numbers were used to label specimens coded in order to preserve the confidentiality of the subjects. Work with these materials was designated as IRB exempt under paragraph 4 and it is covered under IRB protocol #IRB-AAAI7305. GBM-derived GSCs were grown in DMEM:F12 containing 1 \times N2 and B27 supplements (Invitrogen) and human recombinant FGF-2 and EGF (20 ng ml⁻¹ each; Peprotech). Cells at passage (P) 4 were transduced using lentiviral particle in medium containing 4 μ g ml⁻¹ of polybrene (Sigma). Cells were cultured in hypoxic chamber with 1% O₂ (O₂ Control Glove Box, Coy Laboratory Products, MI) for the indicated times or treated with a final concentration of 100–300 μ M CoCl₂ (Sigma) as specified in figure legends.

Mouse neurosphere assay was performed by plating 2,000 cells in 35 mm dishes in collagen containing NSC medium to ensure that distinct colonies were derived from single cells and therefore clonal in origin⁴³. We determined neurosphere formation over serial clonal passages in limiting dilution semi-solid cultures and the cell expansion rate over passages, which is considered a direct indication of self-renewing symmetric cell divisions⁴⁴. For serial sub-culturing we mechanically dissociated neurospheres into single cells in bulk and re-cultured them under the same conditions for six passages. The number of spheres was scored after 14 days. Only colonies >100 μ m in diameter were counted as spheres. Neurosphere size was determined by measuring the diameters of individual neurospheres under light microscopy. Data are presented as percent of neurospheres obtained at each passage (number of neurospheres scored/number of NSCs plated \times 100) in three independent experiments. *P* value was calculated using a multiple *t*-test with Holm–Sidak correction for multiple comparisons. To determine the expansion rate, we plated 10,000 cells from 3 independent P1 clonal assays in 35 mm dishes and scored the number of viable cells after 7 days by Trypan Blue exclusion. Expansion rate of NSCs was determined using a linear regression model and difference in the slopes (*P* value) was determined by the analysis of covariance (ANCOVA) using Prism 6.0 (GraphPad). Limiting dilution assay (LDA) for human GSCs was performed as described previously⁴⁵. Briefly, spheres were dissociated into single cells and plated into 96-well plates in 0.2 ml of medium containing growth factors at increasing densities (1–100 cells per well) in triplicate. Cultures were left undisturbed for

14 days, and then the percent of wells not containing spheres for each cell dilution was calculated and plotted against the number of cells per well. Linear regression lines were plotted, and we estimated the minimal frequency of glioma cells endowed with stem cell capacity (the number of cells required to generate at least one sphere in every well = the stem cell frequency) based on the Poisson distribution and the intersection at the 37% level using Prism 6.0 software. Data represent the means of three independent experiments performed in different days for the evaluation of the effects of ID2, ID2(T27A) in the presence or in the absence of DYRK1B. LDA for the undegradable HIF2 α rescue experiment was performed by using three cultures transduced independently on the same day.

Identification of phosphorylation sites of ID2. To identify the sites of ID2 phosphorylation from IMR32 human neuroblastoma cells, the immunoprecipitated ID2 protein was excised, digested with trypsin, chymotrypsin and Lys-C and the peptides extracted from the polyacrylamide in two 30 μ l aliquots of 50% acetonitrile/5% formic acid. These extracts were combined and evaporated to 25 μ l for MS analysis. The LC–MS system consisted of a state-of-the-art Finnigan LTQ-FT mass spectrometer system with a Protana nanospray ion source interfaced to a self-packed 8 cm \times 75 μ m id Phenomenex Jupiter 10 μ m C18 reversed-phase capillary column. 0.5–5 μ l volumes of the extract were injected and the peptides eluted from the column by an acetonitrile/0.1 M acetic acid gradient at a flow rate of 0.25 μ l min⁻¹. The nanospray ion source was operated at 2.8 kV. The digest was analysed using the double play capability of the instrument acquiring full scan mass spectra to determine peptide molecular weights and product ion spectra to determine amino acid sequence in sequential scans. This mode of analysis produces approximately 1200 CAD spectra of ions ranging in abundance over several orders of magnitude. Tandem MS/MS experiments were performed on each candidate phosphopeptide to verify its sequence and locate the phosphorylation site. A signature of a phosphopeptide is the detection of loss of 98 daltons (the mass of phosphoric acid) in the MS/MS spectrum. With this method, three phosphopeptides were found to carry phosphorylations at residues Ser5, Ser14 and Thr27 of the ID2 protein.

Generation of phospho-ID2-T27 antibody. The anti-phospho-T27-ID2 antibody was generated by immunizing rabbits with a short synthetic peptide containing the phosphorylated T27 (CGISRSK-pT-PVDDPMS) (Yenzym Antibodies, LLC). A two-step purification process was applied. First, antiserum was cross-absorbed against the phospho-peptide matrix to purify antibodies that recognize the phosphorylated peptide. Then, the anti-serum was purified against the un-phosphorylated peptide matrix to remove non-specific antibodies.

Immunoblot, immunoprecipitation and *in vitro* binding assay. Cells were lysed in NP40 lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP40, 1.5 mM Na₃VO₄, 50 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM β -glycerolphosphate and EDTA-free protease inhibitor cocktail (Roche)) or RIPA buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP40, 0.5% sodium dodecyl sulphate, 0.1% sodium dodecyl sulphate, 1.5 mM Na₃VO₄, 50 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM β -glycerolphosphate and EDTA-free protease inhibitor cocktail (Roche)). Lysates were cleared by centrifugation at 15,000 r.p.m. for 15 min at 4 °C. For immunoprecipitation, cell lysates were incubated with primary antibody (hydroxyproline, Abcam, ab37067; VHL, BD, 556347; DYRK1A, Cell Signaling Technology, 2771; DYRK1B, Cell Signaling Technology, 5672) and protein G/A beads (Santa Cruz, sc-2003) or phospho-Tyrosine (P-Tyr-100) Sepharose beads (Cell Signaling Technology, 9419), HA affinity matrix (Roche, 11815016001), Flag M2 affinity gel (Sigma, F2426) at 4 °C overnight. Beads were washed with lysis buffer four times and eluted in 2 \times SDS sample buffer. Protein samples were separated by SDS–PAGE and transferred to polyvinylidene difluoride (PVDF) or nitrocellulose (NC) membrane. Membranes were blocked in TBS with 5% non-fat milk and 0.1% Tween20, and probed with primary antibodies. Antibodies and working concentrations are: ID2 1:500 (C-20, sc-489), GFP 1:1,000 (B-2, sc-9996), HIF2 α /EPAS-1 1:250 (190b, sc-13596), c-MYC (9E10, sc-40), and Elongin B 1:1,000 (FL-118, sc-11447), obtained from Santa Cruz Biotechnology; phospho-Tyrosine 1:1,000 (P-Tyr-100, 9411), HA 1:1,000 (C29F4, 3724), VHL 1:500 (2738), DYRK1A 1:1,000, 2771; DYRK1B 1:1,000, 5672) and RBX1 1:2,000 (D3J5I, 11922), obtained from Cell Signaling Technology; VHL 1:500 (GeneTex, GTX101087); β -actin 1:8000 (A5441), α -tubulin 1:8,000 (T5168), and Flag M2 1:500 (F1804) obtained from Sigma; HIF1 α 1:500 (H1alpha67, NB100-105) and Elongin C 1:1,000 (NB100-78353) obtained from Novus Biologicals; HA 1:1000 (3F10, 12158167001) obtained from Roche. Secondary antibodies horseradish-peroxidase-conjugated were purchased from Pierce and ECL solution (Amersham) was used for detection.

For *in vitro* binding assays, HA-tagged RBX1, Elongin B, Elongin C and VHL were *in vitro* translated using TNT quick coupled transcription/translation system (Promega). Active VHL protein complex was purchased from EMD Millipore. Purified His-VHL protein was purchased from ProteinOne (Rockville, MD).

GST, GST-ID2 and Flag-ID2 proteins were bacterial expressed and purified using glutathione sepharose beads (GE healthcare life science). Active DYRK1B (Invitrogen) was used for *in vitro* phosphorylation of Flag-ID2 proteins. Biotinylated wild-type and modified (pT27 and T27W) ID2 peptides (amino acids 14–34) were synthesized by LifeTein (Somerset, NJ). *In vitro* binding experiments between ID2 and VCB-Cul2 were performed using 500 ng of Flag-ID2 and 500 ng of VCB-Cul2 complex or 500 ng VHL protein in binding buffer (50 mM Tris-Cl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 10 mM β -glycerophosphate, 10 mM sodium pyrophosphate, 50 mM sodium fluoride, 1.5 mM Na_3VO_4 , 0.2% NP40, 10% glycerol, 0.1 mg ml⁻¹ BSA and EDTA-free protease inhibitor cocktail (Roche)) at 4 °C for 3 h. *In vitro* binding between ID2 peptides and purified proteins was performed using 2 μ g of ID2 peptides and 200 ng of recombinant VCB-Cul2 complex or 200 ng recombinant VHL in binding buffer (50 mM Tris-Cl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 10 mM β -glycerophosphate, 10 mM sodium pyrophosphate, 50 mM sodium fluoride, 1.5 mM Na_3VO_4 , 0.4% NP40, 10% glycerol, 0.1 mg ml⁻¹ BSA and EDTA-free protease inhibitor cocktail (Roche)) at 4 °C for 3 h or overnight. Protein complexes were pulled down using glutathione sepharose beads (GE Healthcare Life Science) or streptavidin conjugated beads (Thermo Fisher Scientific) and analysed by immunoblot.

In vitro and in vivo kinase assays. Cdk1, Cdk5, DYRK1A, DYRK1B, ERK, GSK3, PKA, CaMKII, Chk1, Chk2, RSK-1, RSK-2, aurora-A, aurora-B, PLK-1, PLK-2, and NEK2 were all purchased from Life Technology and ATM from EMD Millipore. The 18 protein kinases tested in the survey were selected because they are proline-directed S/T kinases (Cdk1, Cdk5, DYRK1A, DYRK1B, ERK) and/or because they were considered to be candidate kinases for Thr27, Ser14 or Ser5 from kinase consensus prediction algorithms (NetPhosK1.0, <http://www.cbs.dtu.dk/services/NetPhosK/>; GPS Version 3.0 <http://gps.biocuckoo.org/#>) or visual inspection of the flanking regions and review of the literature for consensus kinase phosphorylation motifs. 1 μ g of bacterially purified GST-ID substrates were incubated with 10–20 ng each of the recombinant active kinases. The reaction mixture included 10 μ Ci of [γ -³²P]ATP (PerkinElmer Life Sciences) in 50 μ l of kinase buffer (25 mM Tris-HCl, pH 7.5, 5 mM β -glycerophosphate, 2 mM dithiothreitol (DTT), 0.1 mM Na_3VO_4 , 10 mM MgCl₂, and 0.2 mM ATP). Reactions were incubated at 30 °C for 30 min. Reactions were terminated by addition of Laemmli SDS sample buffer and boiling on 95 °C for 5 min. Proteins were separated on SDS-PAGE gel and phosphorylation of proteins was visualized by autoradiography. Coomassie staining was used to document the amount of substrates included in the kinase reaction. *In vitro* phosphorylation of Flag-ID2 proteins by DYRK1B (Invitrogen) was performed using 500 ng of GST-DYRK1B and 200 ng of bacterially expressed purified Flag-ID2 protein.

In vivo kinase assay in GSCs and glioma cells was performed using endogenous or exogenously expressed DYRK1A and DYRK1B. Cell lysates were prepared in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP40, 1.5 mM Na_3VO_4 , 50 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM β -glycerophosphate and EDTA-free protease inhibitor cocktail (Roche)). DYRK1 kinases were immunoprecipitated using DYRK1A and DYRK1B antibodies (for endogenous DYRK1 proteins) or GFP antibody (for exogenous GFP-DYRK1 proteins) from 1 mg cellular lysates at 4 °C. Immunoprecipitates were washed with lysis buffer four times followed by two washes in kinase buffer as described above and incubated with 200 ng purified Flag-ID2 protein in kinase buffer for 30 min at 30 °C. Kinase reactions were separated by SDS-PAGE and analysed by western blot using p-T27-ID2 antibody.

Protein half-life and stoichiometry. HIF2 α half-life was quantified using ImageJ processing software (NIH). Densitometry values were analysed by Prism 6.0 using the linear regression function. Stoichiometric quantification of ID2 and VHL in U87 cells was obtained using recombinant Flag-ID2 and His-tagged-VHL as references. The chemiluminescent signal of serial dilutions of the recombinant proteins was quantified using ImageJ, plotted to generate a linear standard curve against which the densitometric signal generated by serial dilutions of cellular lysates (1×10^6 U87 cells) was calculated. Triplicate values \pm s.e.m. were used to estimate the ID2:VHL ratio per cell. The stoichiometry of pT27-ID2 phosphorylation was determined as described⁴⁶. Briefly, SK-N-SH cells were plated at density of 1×10^6 in 100 mm dishes. Forty-eight hours later 1.5 mg of cellular lysates from cells untreated or treated with CoCl₂ during the previous 24 h were prepared in RIPA buffer and immunoprecipitated using 4 μ g of pT27-ID2 antibody or rabbit IgG overnight at 4 °C. Immune complexes were collected with TrueBlot anti-rabbit IgG beads (Rockland), washed 5 times in lysis buffer, and eluted in SDS sample buffer. Serial dilutions of cellular lysates, IgG and pT27-ID2 immunoprecipitates were loaded as duplicate series for SDS-PAGE and western blot analysis using ID2 or p-T27-ID2 antibodies. Densitometry quantification of the chemiluminescent signals was used to determine (1) the efficiency of the immunoprecipitation using the antibody against p-ID2-T27 and (2) the ratio between efficiency of the immunoprecipitation evaluated by western blot for p-T27-ID2 and total ID2 antibodies.

This represents the percent of phosphorylated Thr27 of ID2 present in the cell preparation.

Identification of ID2 complexes by mass spectrometry. Cellular ID2 complexes were purified from the cell line NCI-H1299 stably engineered to express Flag-HA-ID2. Cellular lysates were prepared in 50 mM Tris-HCl, 250 mM NaCl, 0.2% NP40, 1 mM EDTA, 10% glycerol, protease and phosphatase inhibitors. Flag-HA-ID2 immunoprecipitates were recovered first with anti-Flag antibody-conjugated M2 agarose (Sigma) and washed with lysis buffer containing 300 mM NaCl and 0.3% NP40. Bound polypeptides were eluted with Flag peptide and further affinity purified by anti-HA antibody-conjugated agarose (Roche). The eluates from the HA beads were analysed directly on long gradient reverse phase LC-MS/MS. A specificity score of proteins interacting with ID2 was computed for each polypeptide by comparing the number of peptides identified from mass spectrometry analysis to those reported in the CRAPome database that includes a list of potential contaminants from affinity purification-mass spectrometry experiments (<http://www.crapome.org>). The specificity score is computed as [(#peptide*#xcorr)/(AveSC*MaxSC*# of Expt.)], #peptide, identified peptide count; #xcorr, the cross-correlation score for all candidate peptides queried from the database; AveSC, averaged spectral counts from CRAPome; MaxSC, maximal spectral counts from CRAPome; and # of Expt., the total found number of experiments from CRAPome.

Ubiquitylation assay. U87 cells were transfected with pcDNA3-HA-HIF α (HIF1 α or HIF2 α), pcDNA3-Flag-ID2 (WT or T27A), pEGFP-DYRK1B and pcDNA3-Myc-Ubiquitin. 36 h after transfection, cells were treated with 20 μ M MG132 (EMD Millipore) for 6 h. After washing with ice-cold PBS twice, cells were lysed in 100 μ l of 50 mM Tris-HCl pH 8.0, 150 mM NaCl (TBS) containing 2% SDS and boiled at 100 °C for 10 min. Lysates were diluted with 900 μ l of TBS containing 1% NP40. Immunoprecipitation was performed using 1 mg of cellular lysates. Ubiquitylated proteins were immunoprecipitated using anti-Myc antibody and analysed by western blot using HA antibody.

Docking of ID2 peptide to the VCB complex. A previously described⁴⁷, highly accurate flexible peptide docking method implemented in ICM software (Molsoft LLC, La Jolla CA) was used to dock ID2 peptides to VCB or components thereof. A series of overlapping peptides of varying lengths were docked to the complex of VHL and Elongin C (EloC), or VHL or EloC alone, from the recent crystallographic structure²² of the VHL-CRL ligase. Briefly, an all-atom model of the peptide was docked into grid potentials derived from the X-ray structure using a stochastic global optimization in internal coordinates with pseudo-Brownian and collective 'probability-biased' random moves as implemented in the ICM program. Five types of potentials for the peptide-receptor interaction energy — hydrogen van der Waals, non-hydrogen van der Waals, hydrogen bonding, hydrophobicity and electrostatics — were precomputed on a rectilinear grid with 0.5 Å spacing that fills a 34 Å \times 34 Å \times 25 Å box containing the VHL-EloC (V-C) complex, to which the peptide was docked by searching its full conformational space within the space of the grid potentials. The preferred docking conformation was identified by the lowest energy conformation in the search. The preferred peptide was identified by its maximal contact surface area with the respective receptor.

ab initio folding and analysis of the peptides was performed as previously described^{48,49}. *ab initio* folding of the ID2 peptide and its phospho-T27 mutant showed that both strongly prefer an α -helical conformation free (unbound) in solution, with the phospho-T27 mutant having a calculated free energy almost 50 kcal-equivalent units lower than the unmodified peptide.

RT-PCR. Total RNA was prepared with Trizol reagent (Invitrogen) and cDNA was synthesized using SuperScript II Reverse Transcriptase (Invitrogen) as described^{42,50}. Semi-quantitative RT-PCR was performed using AccuPrime Taq DNA polymerase (Invitrogen) and the following primers: for HIF2A Fw 5'-GTGCTCCACCGGCCTGTGA-3' and Rv 5'-TTGTCA CACCTATGGCATATACACA-3'; GAPDH Fw 5'-AGAAGGCTGGGGC TCATTTC-3' and Rv 5'-AGGGGCCATCCACAGTCTTC-3'. The quantitative RT-PCR was performed with a Roche480 thermal cycler, using SYBR Green PCR Master Mix from Applied Biosystem.

Primers used in qRT-PCR are: SOX2 Fw 5'-TTGCTGCCTCTTTAA GACTAGGA-3' and Rv 5'-CTGGGGCTCAAACCTCTCTC-3'; NANOG Fw 5'-ATGCCTCACACGAGACTGT-3' and Rv 5'-AAGTGGGTGTGTTTGC CTTTC-3'; POU5F1 Fw 5'-GTGGAGGAAGCTGACAACAA-3' and Rv 5'-ATTCTCCAGGTTGCCTCTCA-3'; FLT1 Fw 5'-AGCCATA AATGGTCTTTGC-3' and Rv 5'-GTGGTTTGCTTGAGCTGTGT-3'; PIK3CA Fw 5'-TGCAAGAATCAGAACAAATGCC-3' and 5'-CAGGAGGCATT CTAAAGTCA-3'; BMI1 Fw 5'-AATCCCCACCTGATGTGTGT-3' and Rv 5'-GCTGGTCTCCAGGTAACGAA-3'; GAPDH Fw 5'-GAAGGTGAAG GTCGGAGTCAAC-3' and Rv 5'-CAGAGTAAAAGCAGCCCTGGT-3'; 18S Fw 5'-CGCCGCTAGAGGTGAAATTC-3' and Rv 5'-CTTTCGC TCTGGTCCGTCTT-3'. The relative amount of specific mRNA was normalized

to 18S or GAPDH. Results are presented as the mean \pm s.d. of three independent experiments each performed in triplicate ($n=9$). Statistical significance was determined by Student's *t*-test (two-tailed) using GraphPad Prism 6.0 software.

Subcutaneous and intracranial xenograft glioma models. Mice were housed in pathogen-free animal facility. All animal studies were approved by the IACUC at Columbia University (numbers AAAE9252; AAAE9956). Mice were 4–6-week-old male athymic nude (*Nu/Nu*, Charles River Laboratories). No statistical method was used to pre-determine sample size. No method of randomization was used to allocate animals to experimental groups. Mice in the same cage were generally part of the same treatment. The investigators were not blinded during outcome assessment. In none of the experiments did tumours exceed the maximum volume allowed according to our IACUC protocol, specifically 20 mm in the maximum diameter. 2×10^5 U87 cells stably expressing a doxycycline inducible lentiviral vector coding for DYRK1B or the empty vector were injected subcutaneously in the right flank in 100 μ l volume of saline solution (7 mice per each group). Mice carrying 150–220 mm³ subcutaneous tumours (21 days after injection) generated by cells transduced with DYRK1B were treated with vehicle or doxycycline by oral gavage (Vibramycin, Pfizer Labs; 8 mg ml⁻¹, 0.2 ml per day)⁵¹; mice carrying tumours generated by cells transduced with the empty vector were also fed with doxycycline. Tumour diameters were measured daily with a caliper and tumour volumes estimated using the formula: width² \times length/2 = *V* (mm³). Mice were euthanized after 5 days of doxycycline treatment. Tumours were dissected and fixed in formalin for immunohistochemical analysis. Data are means \pm s.d. of 7 mice in each group. Statistical significance was determined by ANCOVA using GraphPad Prism 6.0 software package (GraphPad).

Orthotopic implantation of glioma cells was performed as described previously using 5×10^4 U87 cells transduced with pLOC-vector, pLOC-DYRK1B (WT) or pLOC-DYRK1B-K140R mutant in 2 μ l phosphate buffer⁴². In brief, 5 days after lentiviral infection, cells were injected 2 mm lateral and 0.5 mm anterior to the bregma, 2.5 mm below the skull of 4–6-week-old athymic nude (*Nu/Nu*, Charles River Laboratories) mice. Mice were monitored daily for abnormal ill effects according to AAALAS guidelines and euthanized when neurological symptoms were observed. Tumours were dissected and fixed in formalin for immunohistochemical analysis and immunofluorescence using V5 antibody (Life technologies, 46-0705) to identify exogenous DYRK1B and an antibody against human vimentin (Sigma, V6630) to identify human glioma cells. A Kaplan–Meier survival curve was generated using the GraphPad Prism 6.0 software package (GraphPad). Points on the curves indicate glioma related deaths ($n=7$ animals for each group, *p* was determined by log rank analysis). We did not observe non-glioma related deaths. Mice injected with U87 cells transduced with pLOC-DYRK1B(WT) that did not show neurological signs on day 70 were euthanized for histological evaluation and shown as tumour-free mice in Fig. 5g. Intracranial injection of H-Ras-V12-IRES-Cre-ER-shp53 lentivirus was performed in 4-week-old *Id1^{Flox/Flox}, Id2^{Flox/Flox}, Id3^{-/-}* mice (C57Bl6/SV129). Briefly, 1.3 μ l of purified lentiviral particles in PBS were injected 1.45 mm lateral and 1.6 mm anterior to the bregma and 2.3 mm below the skull using a stereotaxic frame. Tamoxifen was administered for 5 days at 9 mg per 40 g of mouse weight by oral gavage starting 30 days after surgery. Mice were killed 2 days later and brains dissected and fixed for histological analysis.

Immunohistochemistry and immunofluorescence. Tissue preparation and immunohistochemistry on tumour xenografts were performed as previously described^{42,50,52}. Antibodies used in immunostaining are: HIF2 α , mouse monoclonal, 1:200 (Novus Biological, NB100-132); Olig2, rabbit polyclonal, 1:200 (IBL International, JP18953); human Vimentin 1:50 (Sigma, V6630), Bromodeoxyuridine, mouse monoclonal 1:500 (Roche, 11170376001), V5 1:500 (Life technologies, 46-0705). Sections were permeabilized in 0.2% tritonX-100 for 10 min, blocked with 1% BSA-5% goat serum in PBS for 1 h. Primary antibodies were incubated at 4 °C overnight. Secondary antibodies biotinylated (Vector Laboratories) or conjugated with Alexa594 (1:500, Molecular Probes) were used. Slides were counterstained with haematoxylin for immunohistochemistry and DNA was counterstained with DAPI (Sigma) for immunofluorescence. Images were acquired using an Olympus IX70 microscope equipped with digital camera and processed using Adobe Photoshop CS6 software. BrdU-positive cells were quantified by scoring the number of positive cells in five 4×10^{-3} mm² images from 5 different mice from each group. Blinding was applied during histological analysis. Data are presented as means of five different mice \pm standard deviation (s.d.) (two-tailed Student's *t*-test, unequal variance).

Computational analysis of dependency of the HIF2 α regulon on ID2 activity. To infer if ID2 modulates the interactions between HIF2 α and its transcriptional targets we used a modified version of MINDY⁵³ algorithm, called CINDY²⁵. CINDY uses adaptive partitioning method to accurately estimate the full conditional mutual information between a transcription factor and a target gene given the expression or activity of a signalling protein. Briefly, for every pair of transcription factor and target gene of interest, it estimates the mutual information that is, how much

information can be inferred about the target gene when the expression of the transcription factor is known, conditioned on the expression/activity of the signalling protein. It estimates this conditional mutual information by estimating the multi-dimensional probability densities after partitioning the sample distribution using adaptive partitioning method. We applied CINDY algorithm on gene expression data for 548 samples obtained from The Cancer Genome Atlas (TCGA). Since the activity level and not the gene expression of ID2 is the determinant of its modulatory function that is, the extent to which it modulates the transcriptional network of HIF2 α , we used an algorithm called Virtual Inference of Protein-activity by Enriched Regulon analysis (VIPER) to infer the activity of ID2 protein from its gene expression profile²⁶. VIPER method allows the computational inference of protein activity, on an individual sample basis, from gene expression profile data. It uses the expression of genes that are most directly regulated by a given protein, such as the targets of a transcription factor (TF), as an accurate reporter of its activity. We defined the targets of ID2 by running ARACNe algorithm on 548 gene expression profiles and use the inferred 106 targets to determine its activity (Supplementary Table 3).

We applied CINDY on 277 targets of HIF2 α represented in Ingenuity pathway analysis (IPA) and for which gene expression data was available (Supplementary Table 4). Of these 277 targets, 77 are significantly modulated by ID2 activity (*P* value ≤ 0.05). Among the set of target genes whose expression was significantly positively correlated (*P* value ≤ 0.05) with the expression of HIF2 α irrespective of the activity of ID2, that is, correlation was significant for samples with both high and low activity of ID2, the average expression of target genes for a given expression of HIF2 α was higher when the activity of ID2 was high. The same set of target gene were more correlated in high ID2 activity samples compared to any set of random genes of same size (Fig. 5a), whereas they were not in ID2 low activity samples (Fig. 5b). We selected 25% of all samples with the highest/lowest ID2 activity to calculate the correlation between HIF2 α and its targets.

To determine whether regulation of ID2 by hypoxia might impact the correlation between high ID2 activity and HIF2 α shown in Fig. 5a, b we compared the effects of ID2 activity versus ID2 expression for the transcriptional connection between HIF2 α and its targets. We selected 25% of all patients ($n=548$) in TCGA with high ID2 activity and 25% of patients with low ID2 activity and tested the enrichment of significantly positively correlated targets of HIF2 α in each of the groups. This resulted in significant enrichment (*P* value < 0.001) in high ID2 activity but showed no significant enrichment (*P* value = 0.093) in low ID2 activity samples. Moreover, the difference in the enrichment score (Δ ES) in these two groups was statistically significant (*P* value < 0.05). This significance is calculated by randomly selecting the same number of genes as the positively correlated targets of HIF2 α , and calculating the Δ ES for these randomly selected genes, giving Δ ES_{rand}. We repeated this step 1,000 times to obtain 1,000 Δ ES_{rand} that are used to build the null distribution (Extended Data Fig. 9b). We used the null distribution to estimate *P* value calculated as (number of Δ ES $>$ Δ ES_{rand})/1,000. Enrichment was observed only when ID2 activity was high but not when ID2 activity was low, thus suggesting that ID2 activity directionally impacts the regulation of targets of HIF2 α by HIF2 α . Consistently, the significant Δ ES using ID2 activity suggests that ID2 activity is determinant of correlation between HIF2 α and its targets.

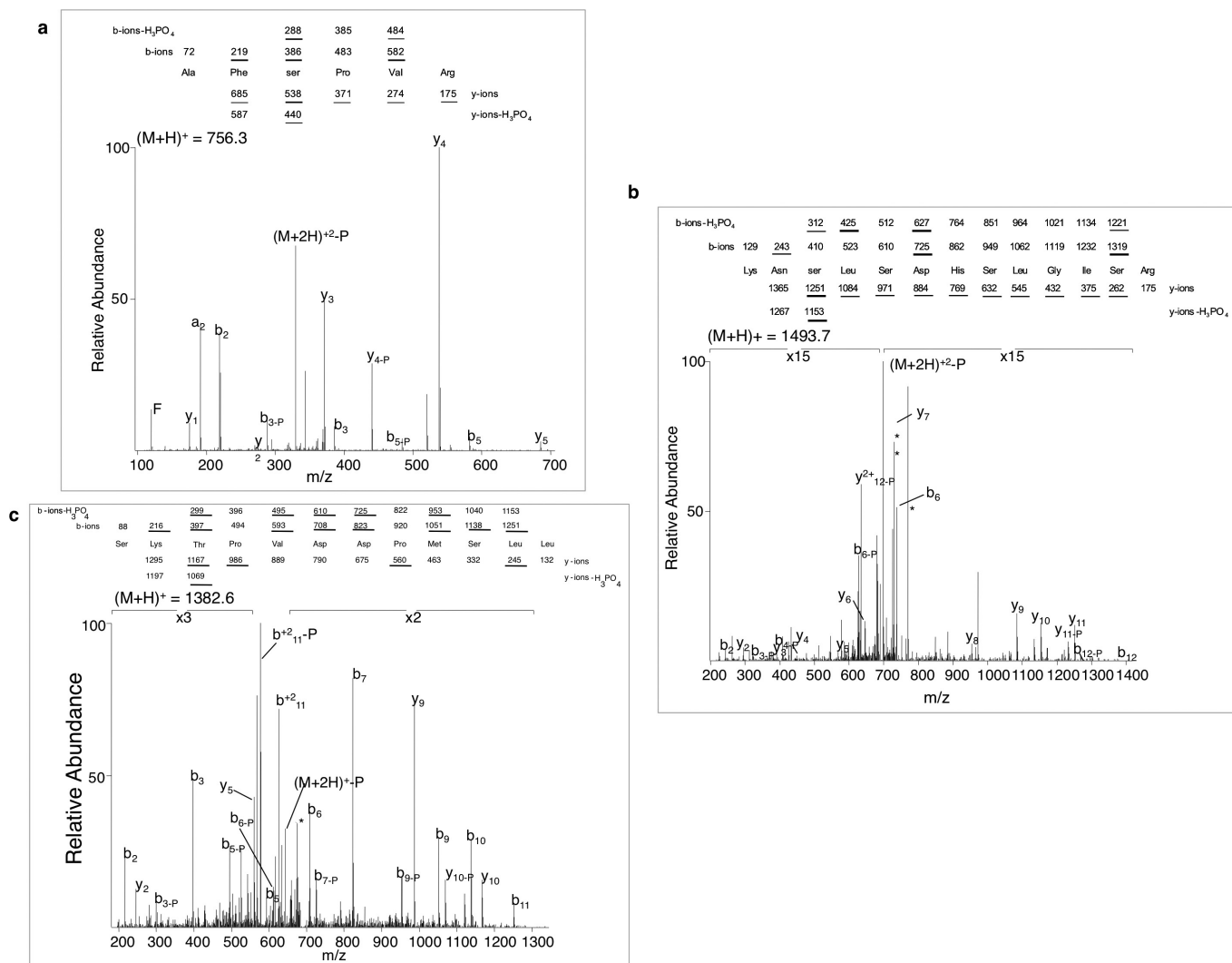
Conversely, when we performed similar analysis using ID2 expression instead of ID2 activity, we found significant enrichment of positively correlated targets of HIF2 α both in samples with high expression (*P* value = 0.025) and low expression of ID2 (*P* value = 0.048). Given the significant enrichment in both groups, we did not observe any significant difference in the enrichment score in the two groups (*P* value of Δ ES = 0.338). Thus, while the determination of the ID2 activity and its effects upon the HIF2 α -targets connection by VIPER and CINDY allowed us to determine the unidirectional positive link between high ID2 activity and HIF2 α transcription, a similar analysis performed using ID2 expression contemplates the dual connection between ID2 and HIF2 α .

Kaplan–Meier analysis for DYRK1A and DYRK1B in human GBM. To test if expression of DYRK1A and DYRK1B is a predictor of prognosis, we divided the patients into two cohorts based on their relative expression compared to the mean expression of all patients in GBM. First cohort contained the patients with high expression of both DYRK1A and DYRK1B ($n=101$) and the other cohort contained patients with low expression ($n=128$). We used average expression for both DYRK1A and DYRK1B, which individually divide the patient cohort into half and half. However, when we use the condition that patients should display higher or lower average expression of both these genes, then we select approximately 19% for high expression and 24% for low expression. Selection of these patients was entirely dependent on the overall expression of these genes in the entire cohort rather than a predefined cutoff. Kaplan–Meier survival analysis showed the significant survival benefit for the patients having the high expression of both DYRK1A and DYRK1B (*P* value = 0.004) compared to the patients with low expression.

When similar analysis was performed using only the expression of DYRK1A or DYRK1B alone, the prediction was either non-significant (DYRK1A) or less significant (DYRK1B, P value = 0.008) when compared to the predictions using the expression of both genes.

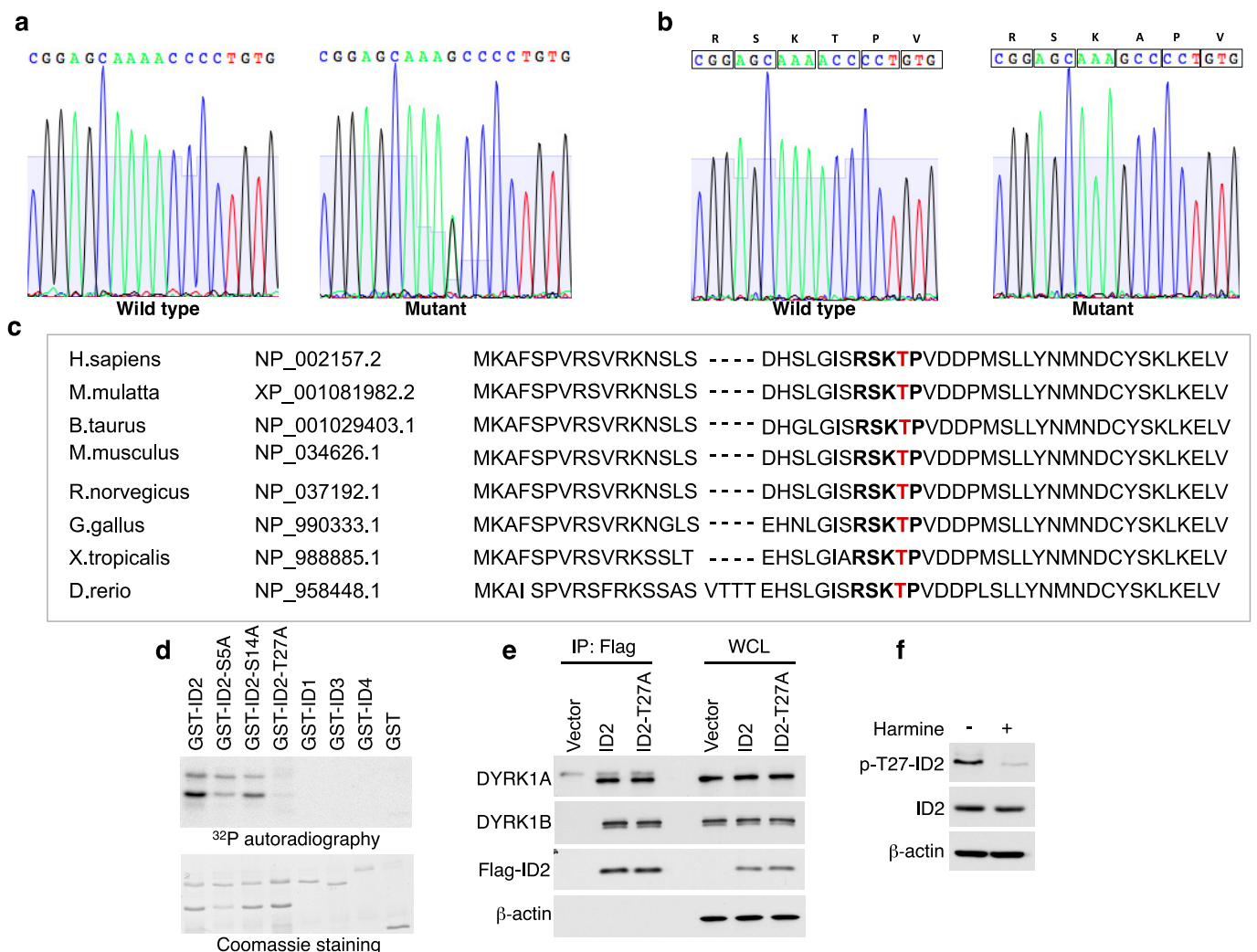
Statistics. Results in graphs are expressed as means \pm s.d. or means \pm s.e.m., as indicated in figure legends, for the indicated number of observations. Statistical significance was determined by the Student's t -test (two-tailed, unequal variance). P value < 0.05 is considered significant and is indicated in figure legends.

42. Carro, M. S. *et al.* The transcriptional network for mesenchymal transformation of brain tumours. *Nature* **463**, 318–325 (2010).
43. Reynolds, B. A. & Weiss, S. Clonal and population analyses demonstrate that an EGF-responsive mammalian embryonic CNS precursor is a stem cell. *Dev. Biol.* **175**, 1–13 (1996).
44. Deleyrolle, L. P. *et al.* Determination of somatic and cancer stem cell self-renewing symmetric division rate using sphere assays. *PLoS One*. **6**, e15844 (2011).
45. Tropepe, V. *et al.* Distinct neural stem cells proliferate in response to EGF and FGF in the developing mouse telencephalon. *Dev. Biol.* **208**, 166–188 (1999).
46. Miinea, C. P. & Lienhard, G. E. Stoichiometry of site-specific protein phosphorylation estimated with phosphopeptide-specific antibodies. *Biotechniques* **34**, 828–831 (2003).
47. Bordner, A. J. & Abagyan, R. Ab initio prediction of peptide-MHC binding geometry for diverse class I MHC allotypes. *Proteins* **63**, 512–526 (2006).
48. Abagyan, R. & Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **235**, 983–1002 (1994).
49. Almond, D. & Cardozo, T. Assessment of immunologically relevant dynamic tertiary structural features of the HIV-1 V3 loop crown R2 sequence by ab initio folding. *J. Vis. Exp.* **43**, 2118 (2010).
50. Zhao, X. *et al.* The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. *Nature Cell Biol.* **10**, 643–653 (2008).
51. Cawthorne, C., Swindell, R., Stratford, I. J., Dive, C. & Welman, A. Comparison of doxycycline delivery methods for Tet-inducible gene expression in a subcutaneous xenograft model. *J. Biomol. Tech.* **18**, 120–123 (2007).
52. Zhao, X. *et al.* The N-Myc-DLL3 cascade is suppressed by the ubiquitin ligase Huwe1 to inhibit proliferation and promote neurogenesis in the developing brain. *Dev. Cell* **17**, 210–221 (2009).
53. Wang, K. *et al.* Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nature Biotechnol.* **27**, 829–837 (2009).



Extended Data Figure 1 | ID2 is phosphorylated on Ser5, Ser14 and Thr27. Chromatographic results of mass spectrometry analysis of ID2 protein immunoprecipitated from IMR32 human neuroblastoma cells.

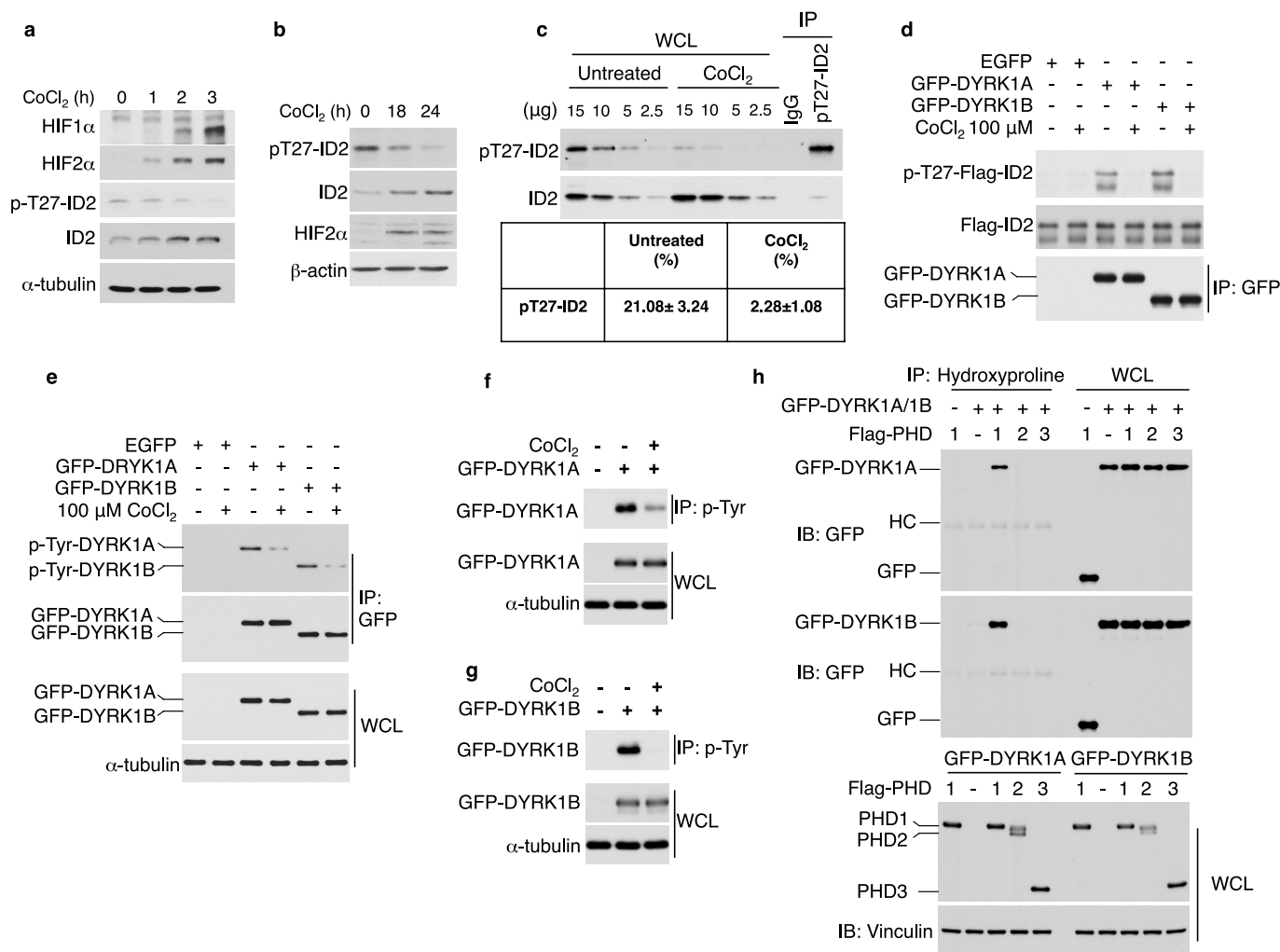
a, The peptide identified as ID2 A3–R8 shows phosphorylation of Ser5.
b, The peptide identified as ID2 K12–R24 shows phosphorylation of Ser14.
c, The peptide identified as ID2 S25–L36 shows phosphorylation of Thr27.



Extended Data Figure 2 | T27A missense mutation in ID2 in human cancer cells and Thr27 phosphorylation of ID2 by DYRK1 kinases.

a, Sequence analysis of genomic DNA from the neuroblastoma cell line IMR32 shows the wild-type sequence (left). Sequencing of DNA from the colon cancer cell line HRT-18 shows a heterozygous mutation resulting in the change of codon-27 from ACC (Thr) to GCC (Ala) (right). **b**, Both wild-type and mutant ID2(T27A) are expressed in HRT-18 colon cancer cells. Sequence analysis of representative clones (out of 20 clones) derived from HRT-18 cDNA demonstrates expression of wild-type (left panel) and mutant (right panel) alleles. **c**, Amino acid sequence flanking Thr27

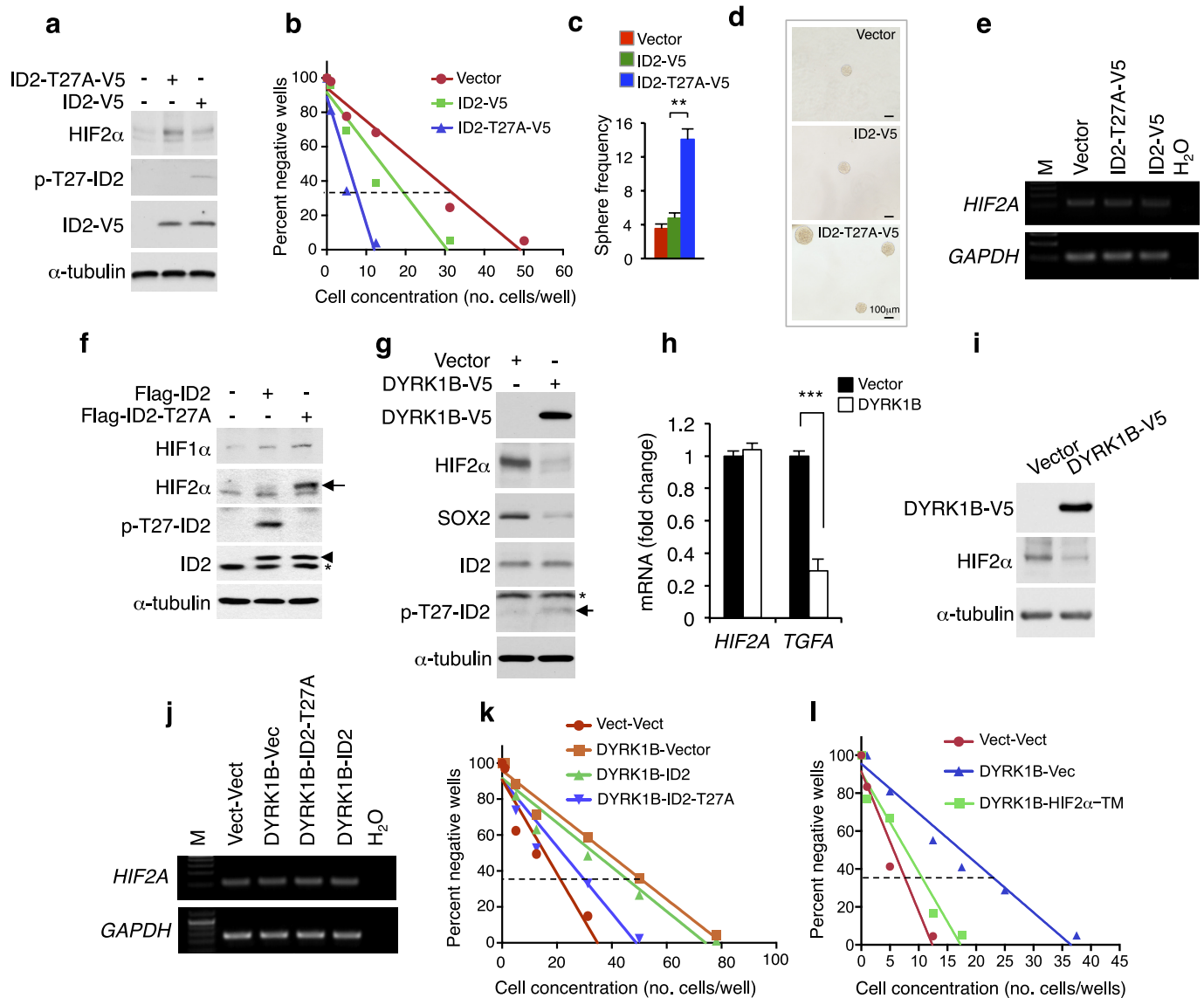
of ID2 (marked in red), including the DYRK1 consensus motif (bold), is evolutionarily conserved. **d**, *In vitro* kinase assay using bacterially expressed GST-ID proteins and recombinant DYRK1A. **e**, U87 cells transfected with Flag-ID2, Flag-ID2(T27A) or the empty vector were immunoprecipitated with Flag antibody. Co-precipitated proteins were analysed by western blot using DYRK1A, DYRK1B and Flag antibodies. β-actin was used as control for loading. WCL, whole cellular lysate. **f**, U87 stably transfected with Flag-ID2 were treated with harmine (10 μM) or vehicle for 24 h and analysed by western blot using the indicated antibodies.



Extended Data Figure 3 | DYRK1 kinase activity and Thr27

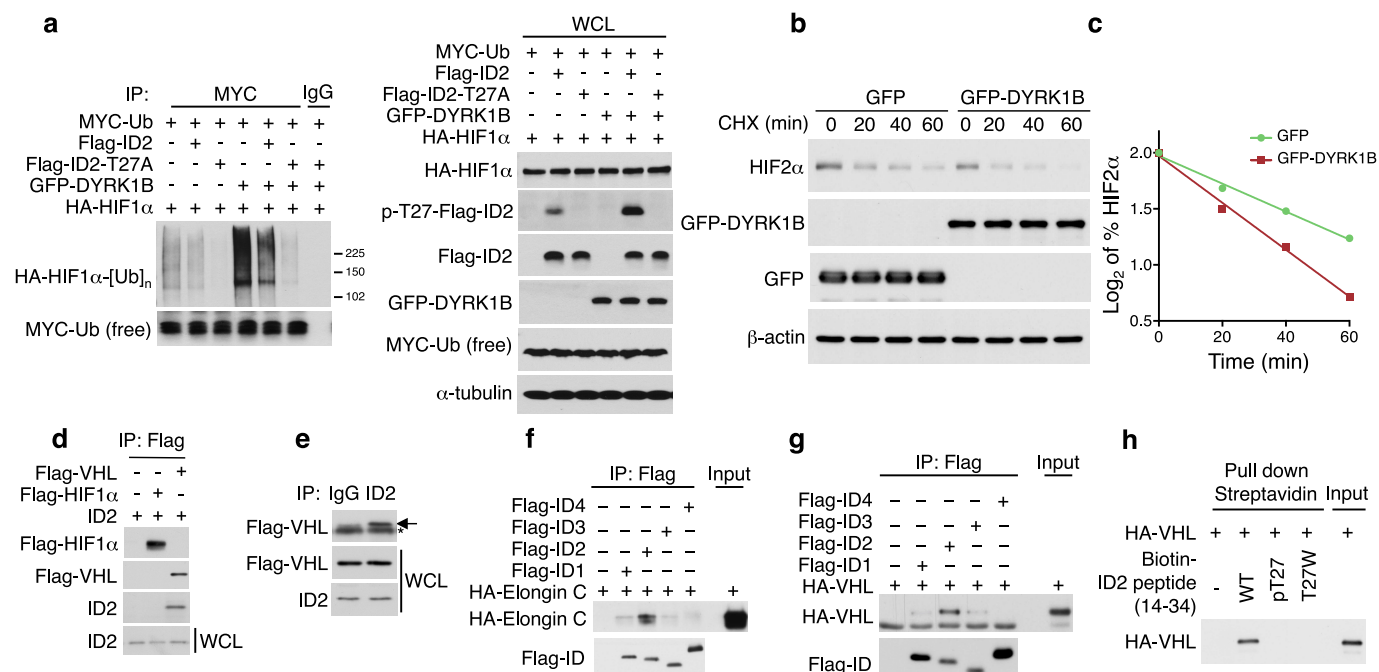
phosphorylation of ID2 are inhibited by hypoxia. **a**, U87 glioma cells were treated with 100 μM CoCl₂ for the indicated times. Cellular lysates were analysed by western blot using the indicated antibodies. **b**, SK-N-SH cells were treated with 300 μM CoCl₂ for the indicated times and assayed by western blot using the indicated antibodies. **c**, Stoichiometric evaluation of pThr-27-ID2 in SK-N-SH cells untreated or treated with CoCl₂ for 24 h. Cellular lysates prepared in denaturing buffer were immunoprecipitated using pT27-ID2 antibody or normal rabbit IgG. Aliquots of whole cellular lysates (WCL, μg) and immunoprecipitates were assayed by western blot using pT27-ID2 and non-phosphorylated ID2 antibodies (upper panels). The efficiency of immunoprecipitation with anti-pT27-ID2 antibody from untreated cells was determined to calculate the percent of the pT27-ID2 in the absence and in the presence of CoCl₂ (lower panel). **d**, 293T cells expressing GFP-DYRK1 proteins untreated or treated with 100 μM CoCl₂ for 12 h were used as a source of active kinase. The kinase activity of the anti-GFP-DYRK1 immunoprecipitates was tested *in vitro* using bacterially expressed and purified Flag-ID2 as substrate. Kinase reactions were evaluated by western blot using p-T27-ID2 antibodies (top). Analysis of

kinase reactions by Flag immunoblot shows similar amount of ID2 protein in each kinase reaction (middle). Immunocomplexes were analysed by western blot using GFP antibody (bottom). **e**, Lysates from U251 cells expressing GFP-DYRK1 proteins untreated or treated with 100 μM CoCl₂ for 6 h were immunoprecipitated using GFP antibodies. Western blot was performed using anti-p-Tyrosine (p-Tyr) or GFP antibodies. Analysis of WCL shows similar expression levels of DYRK1 proteins. α-tubulin was used as control for loading. **f**, Lysates from 293T cells expressing GFP-DYRK1A untreated or treated with 100 μM CoCl₂ for 12 h were immunoprecipitated with anti-p-Tyr antibodies and analysed by western blot using antibodies against GFP. α-tubulin was used as control for loading. **g**, Lysates from 293T cells expressing GFP-DYRK1B untreated or treated with 100 μM CoCl₂ for 12 h were immunoprecipitated with anti-p-Tyr antibodies and analysed by western blot using antibodies against GFP. α-tubulin was used as control for loading. **h**, U87 transfected with GFP-DYRK1A, GFP-DYRK1B or GFP and Flag-PHD1, Flag-PHD2, or Flag-PHD3 were immunoprecipitated using anti-hydroxyproline antibody. Western blot was performed using GFP antibody (upper panels). HC, IgG heavy chain. Lower panels, WCL.



Extended Data Figure 4 | The DYRK1-ID2 Thr27 pathway controls GSCs and HIF2α. **a**, GSC line 48 cells were transduced with lentiviruses expressing ID2 (WT), ID2 (T27A), or the empty vector. **b**, Cells were analysed by *in vitro* LDA. Representative regression plot used to calculate gliomasphere frequency in panel c. **c**, The frequency of cells capable of forming gliospheres by *in vitro* LDA. Data in the histograms represent means of 3 biological replicates \pm s.d.; $**P = 0.00163$. **d**, The microphotographs show representative gliosphere cultures of cells treated as in **a**. **e**, HIF2α mRNAs from cells treated as in **a** were analysed by semi-quantitative RT-PCR. **f**, U87 cells stably expressing Flag-ID2 or Flag-ID2 (T27A) were analysed by western blot using the indicated antibodies. Arrow points to specific band. Arrowhead indicates Flag-ID2. Asterisk indicates endogenous ID2. **g**, GSC line 34 cells were transduced with lentiviruses expressing DYRK1B-V5 or empty vector. Cells were analysed by western blot using the indicated antibodies. Arrow points to specific band. Asterisk indicates a non-specific band. **h**, qRT-PCR from

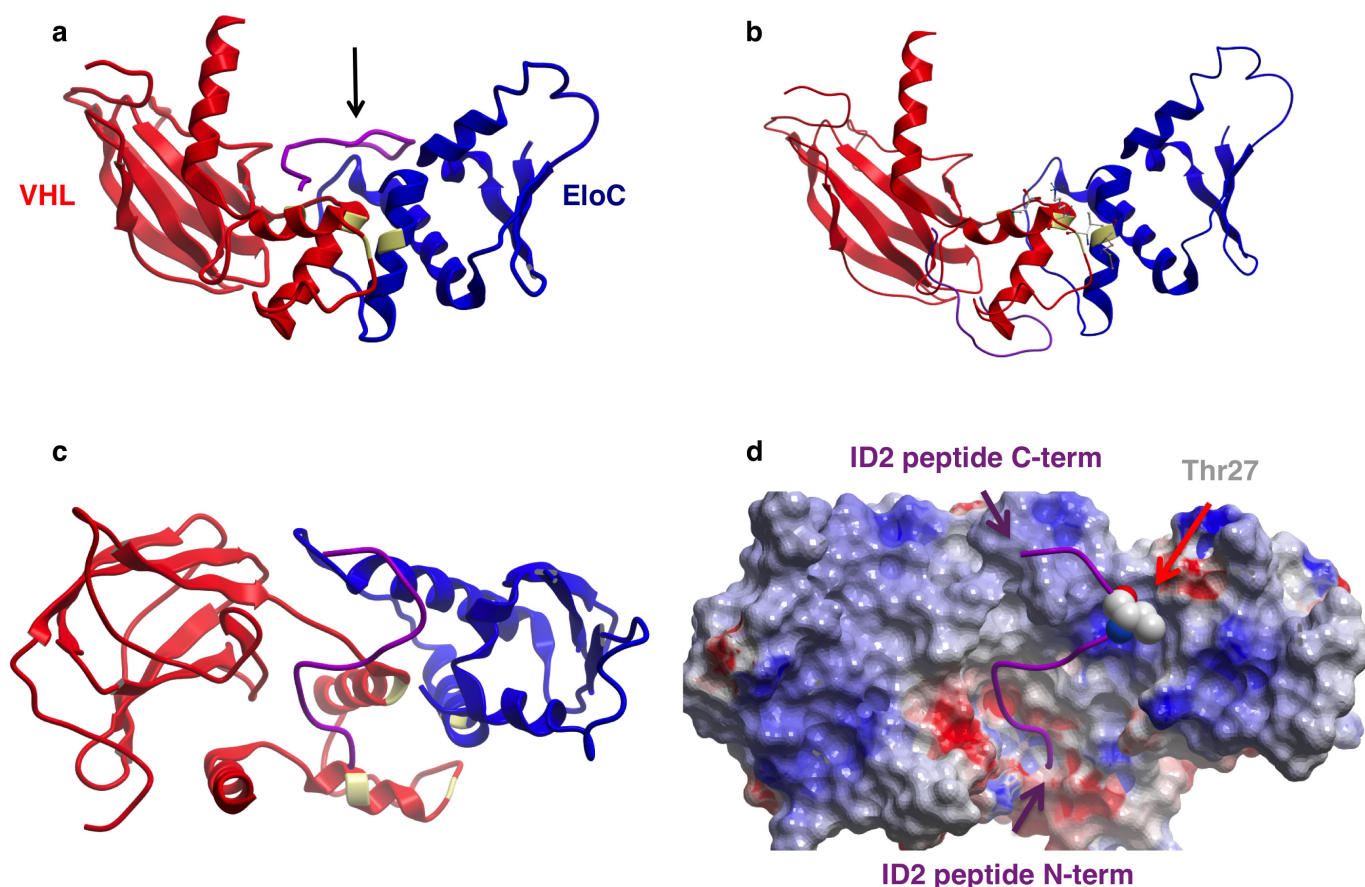
cells treated as in **g**. Data in the histograms represent means \pm s.d. ($n = 9$, triplicate experiments each performed in triplicate; $***P = 8.44524 \times 10^{-7}$ for TGFA). **i**, GSC line 31 was transduced with lentiviruses expressing DYRK1B-V5 or empty vector. Expression of HIF2α, DYRK1B-V5 and α-tubulin was analysed by western blot. **j**, mRNAs from experiment shown in Fig. 3a-c were analysed by semi-quantitative RT-PCR for HIF2α. **k**, GSC line 31 cells were transduced with lentiviruses expressing DYRK1B and ID2, ID2 (T27A), or the empty vector. Cells were analysed by LDA. Representative regression plot used to calculate gliomasphere frequency in Fig. 3b. **l**, GSC cell line 31 cells were transduced with lentiviruses expressing DYRK1B or the empty vector in the absence or in the presence of undegradable HIF2α (HIF2α-TM). Cells were analysed by *in vitro* LDA. Representative regression plot used to calculate the frequency of gliospheres in cultures from three independent infections (Vect plus Vect = 13.55%; DYRK1B-Vect = 4.36%; DYRK1B-HIF2α-TM = 9.73%).



Extended Data Figure 5 | The DYRK1-ID2-Thr27 pathway modulates HIF α stability by regulating the interaction between ID2 and VHL.

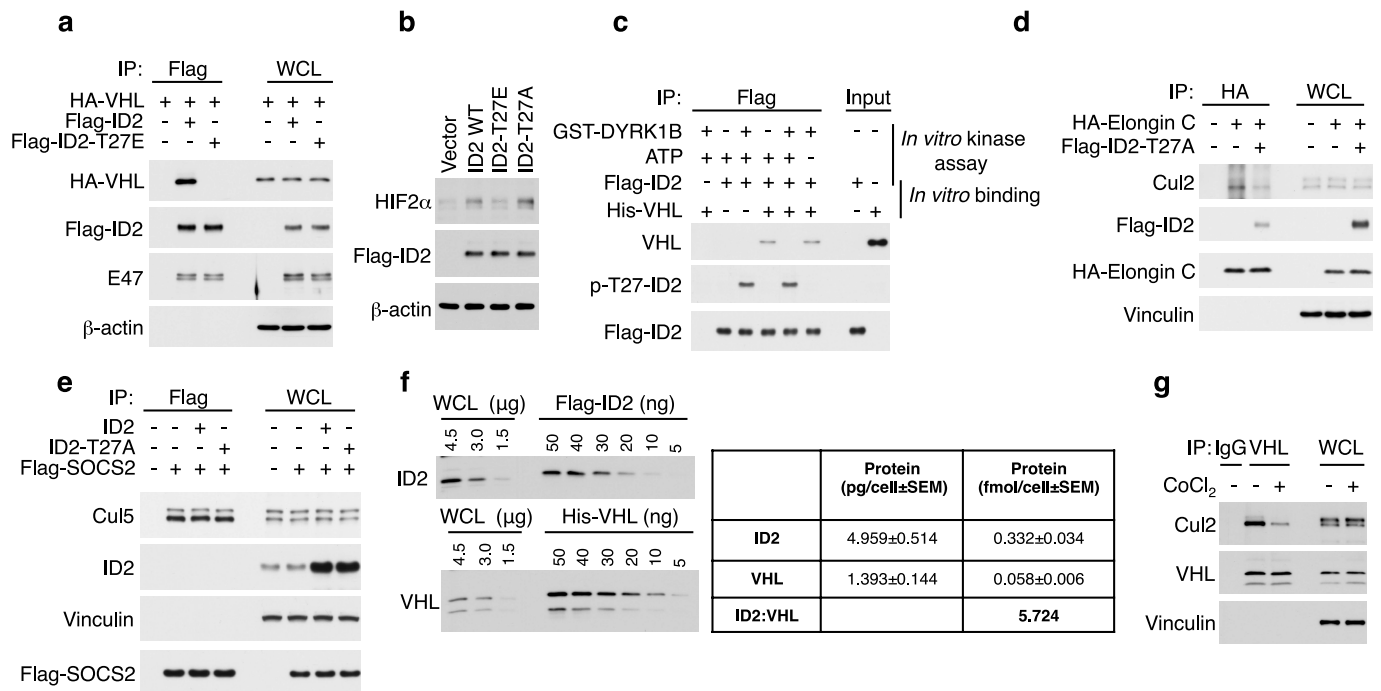
a, *In vivo* ubiquitylation of HIF1 α protein. U87 cells transfected with the expression plasmids HIF1 α and MYC-ubiquitin were co-transfected with Flag-ID2, Flag-ID2(T27A), or the empty vector in the presence or in the absence of GFP-DYRK1B. After treatment with MG132 (20 μ M) for 6 h, lysates were prepared in denaturing buffer and identical aliquots were immunoprecipitated with antibodies directed against MYC. An anti-HA antibody was used to detect HIF1 α ubiquitin conjugates (left); Cellular lysates, WCL, were analysed by western blot using the indicated antibodies (right). **b**, U87 cells were co-transfected with plasmids expressing HA-HIF2 α and GFP-DYRK1B or GFP-vector. Cells were treated with 50 μ g ml $^{-1}$ of CHX for the indicated times and analysed by western blot. **c**, Quantification of HIF2 α protein from the experiment in panel **b** as the log $_2$ of the percent of HIF2 α relative to untreated cells.

d, IMR32 cells were co-transfected with ID2 and Flag-VHL or Flag-HIF1 α expression vectors. Immunoprecipitation was performed using Flag antibody and immunocomplexes and whole cellular lysates (WCL) were analysed by western blot using the indicated antibodies. **e**, IMR32 cells transfected with Flag-VHL expression vector were used for IgG or ID2 antibody immunoprecipitation. Immunocomplexes and WCL were analysed by western blot. Arrow points to the specific Flag-VHL band; asterisk indicates IgG light chain. **f**, Flag immunoprecipitation of binding reactions of *in vitro* translated Flag-ID and HA-Elongin C proteins. Immunocomplexes were analysed by western blot for HA and Flag. **g**, Flag-ID proteins and HA-VHL were translated and incubated *in vitro*. Flag immunocomplexes were analysed by western blot for HA and Flag. **h**, *In vitro* streptavidin pulldown assay of biotinylated ID2 peptides (amino acid 14–34 (WT), pT27, and T27W) and *in vitro* translated HA-VHL. Bound polypeptides were detected by western blot.



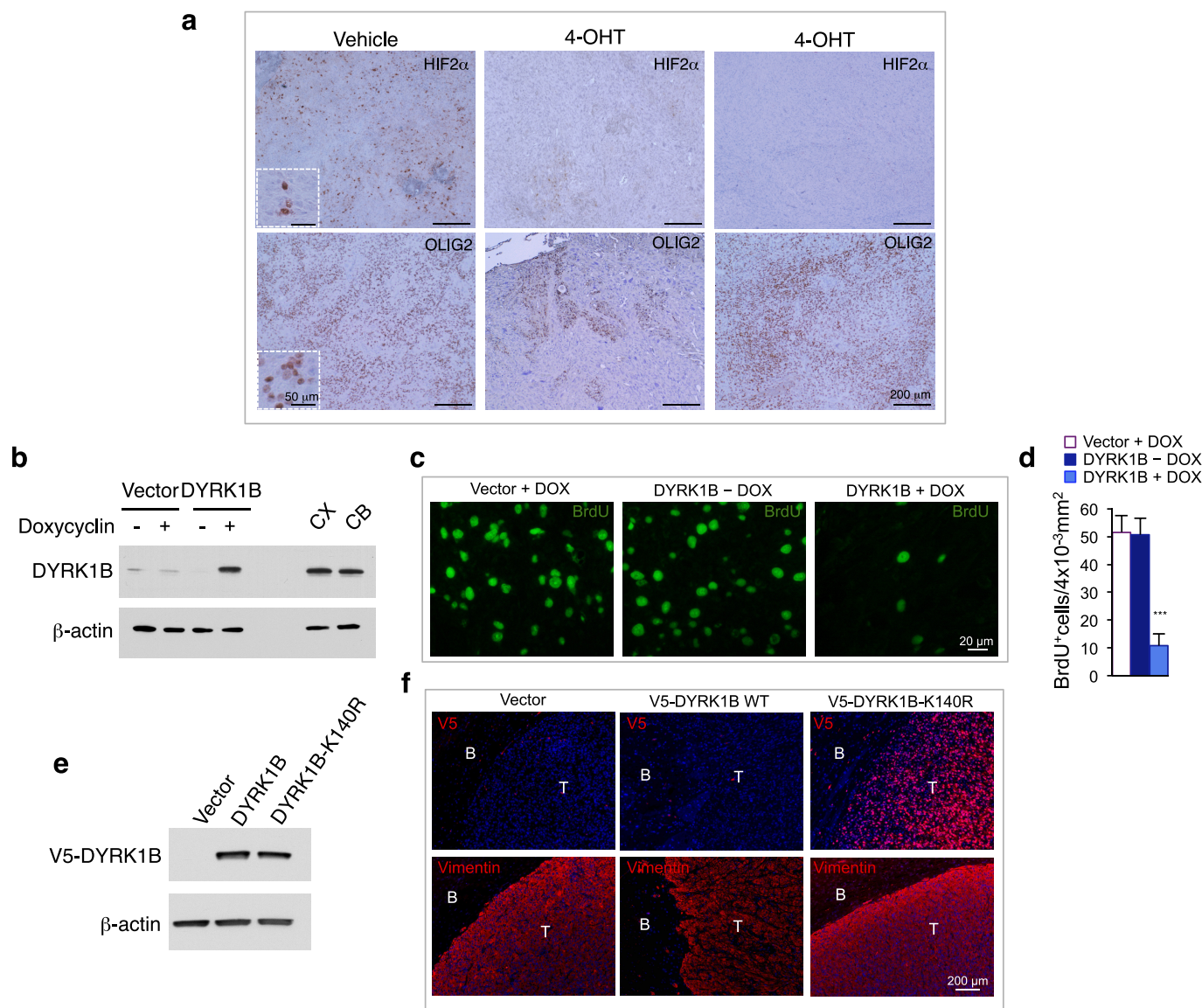
Extended Data Figure 6 | Molecular docking of an ID2 (15–31) peptide on the VHL–Elongin C complex. **a**, Ribbon representation of the backbone of the VHL–Elongin C complex and the predicted binding conformation of the ID2 peptide. VHL (red ribbon), Elongin C (blue ribbon) and the docked ID2 peptide (purple ribbon). Cul2 contact residues are colored yellow ribbon in both VHL and Elongin C. Arrow indicates the ID2 peptide. **b**, Docking result for the phospho-Thr-27-ID2 peptide shown from the same perspective as in panel **a**. **c**, The view and

complex in **a** is rotated 90 degrees around an axis parallel to the page so that the perspective is from the arrow shown in panel **a**. **d**, Electrostatic molecular surface representation of the VHL–Elongin C complex with the docked ID2 peptide. The perspective is the same as in panel **c**. The T27 side chain is shown as space-filling spheres and is indicated by the red arrow. The N-terminus and C-terminus of the ID2 peptide are indicated by purple arrows.



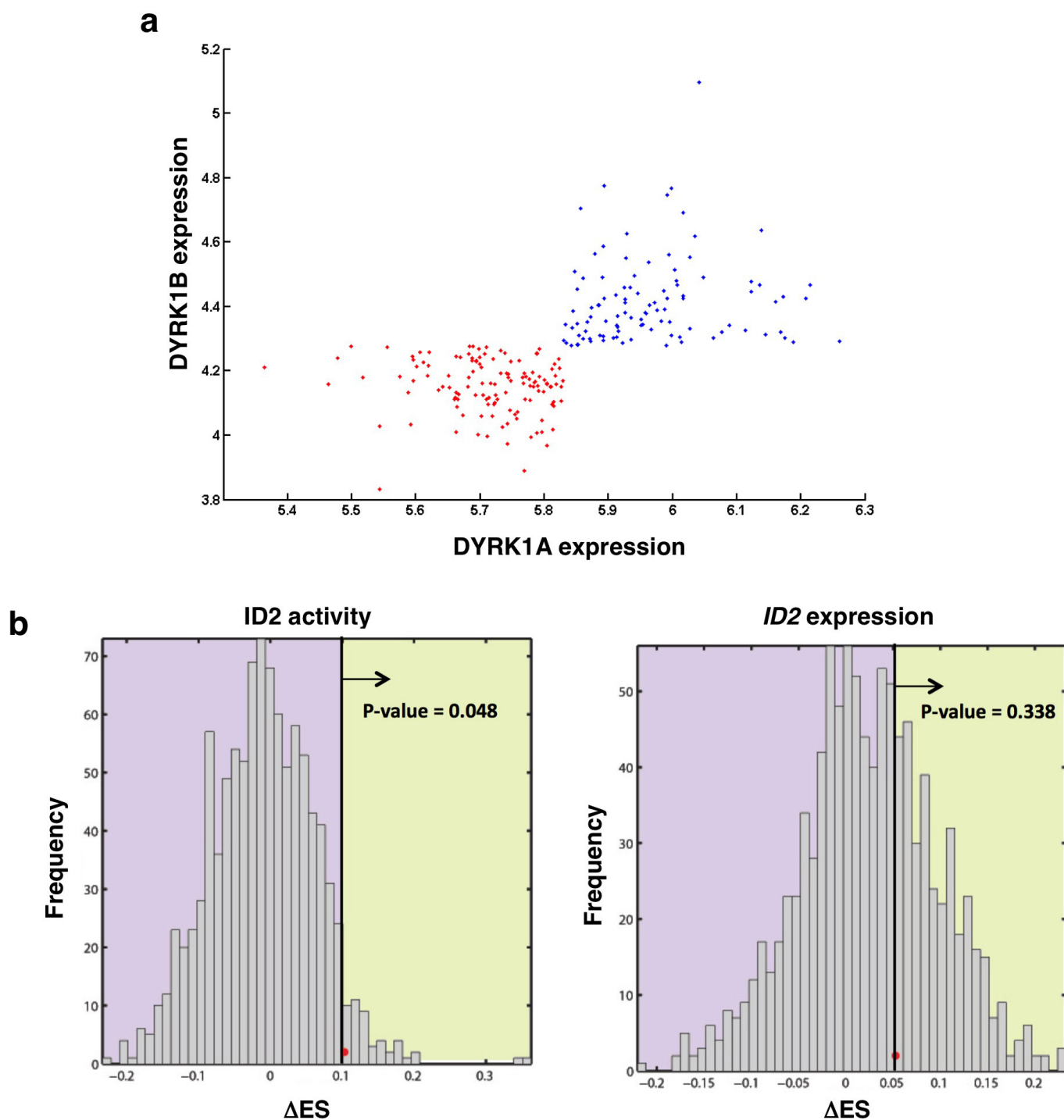
Extended Data Figure 7 | DYRK1-mediated phosphorylation of ID2 prevents dissociation of the VCB-Cul2 complex. a, *In vivo* binding assay using lysates from U87 cells co-transfected with HA-VHL and Flag-ID2 or Flag-ID2(T27E) expression vectors. Flag immunocomplexes were analysed by western blot using HA and Flag antibodies. Whole cell lysates, WCL, were analysed by western blot using the indicated antibodies. Binding of Flag-ID2 and Flag-ID2(T27E) to the bHLH protein E47 is shown as control for ID2 binding. **b**, U87 cells were transfected with Flag-ID2, Flag-ID2(T27A) or Flag-ID2(T27E) plasmids. Cellular lysates were analysed by western blot using the indicated antibodies. **c**, *In vitro* binding between purified Flag-ID2 and His-VHL following *in vitro* kinase reaction using recombinant DYRK1B and Flag-ID2. **d**, Analysis of the HA-Elongin C immunocomplexes in U87 cells transfected with HA-Elongin C in the

absence or presence of Flag-ID2(T27A). Anti-HA immunoprecipitation reactions and WCL were analysed by western blot using antibodies against Cul2, HA (Elongin C), and Flag (ID2). **e**, Analysis of the Flag-SOCS2 immunocomplexes in U87 cells transfected with ID2, ID2(T27A) or the empty vector. Flag immunoprecipitation reactions and WCL were analysed by western blot using antibodies against Cul5, ID2, and Flag (SOCS2). **f**, Stoichiometric analysis of ID2 and VHL in cellular lysates. Decreasing amount of WCL from 1×10^6 U87 cells and purified proteins were assayed by western blot (left). Regression plots of densitometry analysis were used to determine ID2 and VHL protein concentration and the ID2:VHL ratio (right). **g**, Immunoprecipitation of endogenous VHL in U87 cells in the presence and in the absence of CoCl₂. Western blot for Cul2 and VHL are analysed by western blot. Vinculin is shown as loading control.



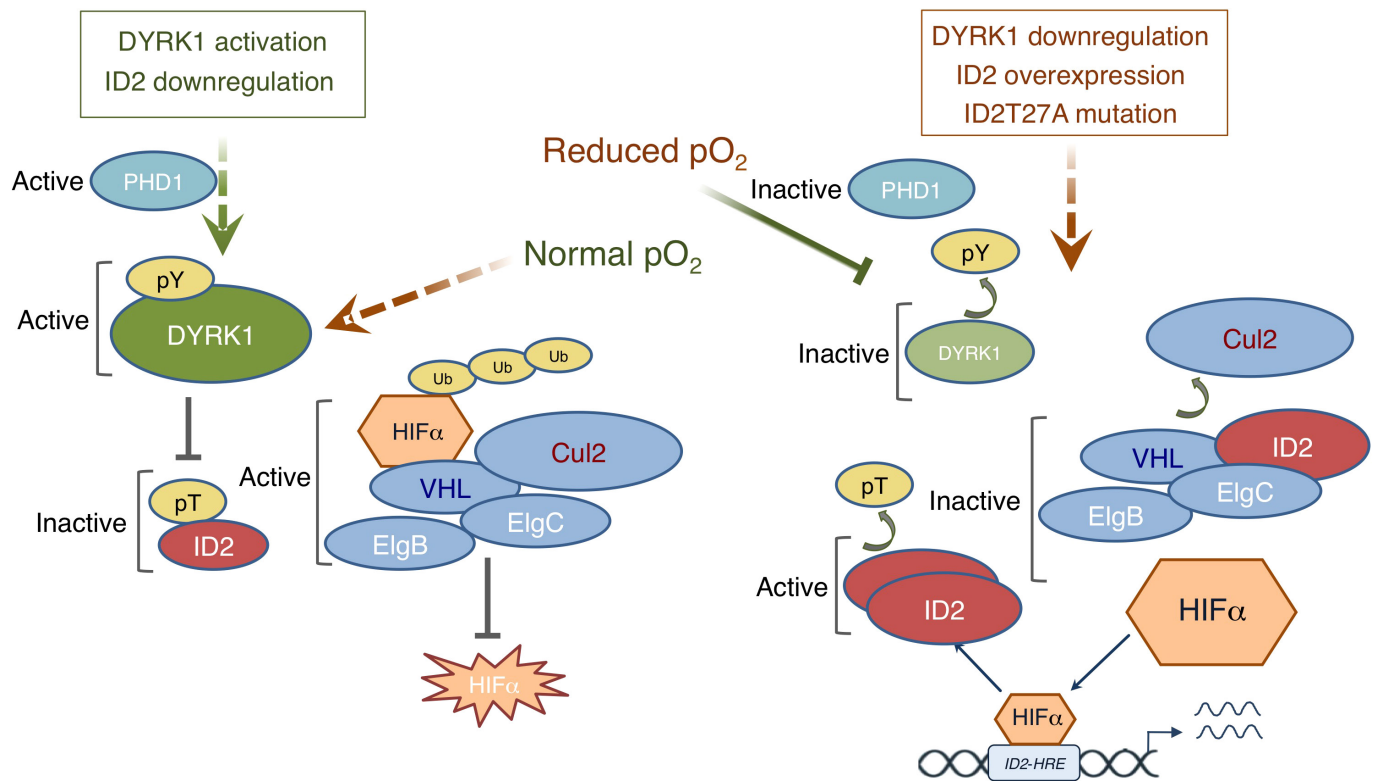
Extended Data Figure 8 | DYRK1 kinase inhibits proliferation of human glioma. **a**, Malignant glioma were induced in *Id1^{Flox/Flox} Id2^{Flox/Flox} Id3^{-/-}* mice via injection of lentivirus expressing H-RAS-V12-IRES-CRE-ER linked to U6-shp53 cassette into the dentate gyrus as described⁹. Mice were treated for 5 days with tamoxifen or vehicle and euthanized 2 days later. Tumours were analysed by immunohistochemistry using HIF2 α and OLIG2 antibodies. Nuclei were counterstained with haematoxylin. **b**, Western blot analysis of DYRK1B in U87 cells stably expressing a doxycycline inducible DYRK1B or the empty vector. Cells were treated with 0.75 $\mu\text{g ml}^{-1}$ doxycycline or vehicle for 36 h. Lysates of adult mouse cortex (CX) and cerebellum (CB) were used to compare exogenous DYRK1B with endogenous levels of the protein. **c**, Tissue

sections from experiment in Fig. 5e, f were analysed by immunostaining using BrdU antibodies. **d**, Quantification of BrdU positive cells from the experiment in **c**. Data in the histograms represent means \pm s.d. ($n = 5$; $***P = 3.065 \times 10^{-7}$, DYRK1B - Dox versus DYRK1B + Dox). Asterisks indicate statistical significance by two-tailed *t*-test. **e**, Western blot analysis of ectopically expressed V5-DYRK1B, V5-DYRK1B-K140R in U87 cells. **f**, Brain cross-sections of mice intracranially injected with U87 cells in **e** were analysed by immunofluorescence using V5 antibody (red, upper panels) to identify exogenous DYRK1B and human vimentin antibody (red, lower panels) to identify human glioma cells. Nuclei were counterstained with DAPI (blue). T, tumour; B, brain.



Extended Data Figure 9 | Analysis of DYRK1A, DYRK1B and ID2 expression in human GBM. **a**, Scatter plot showing the expression of DYRK1A and DYRK1B in GBM. Blue and red dots indicate GBM samples with high or low expression of both DYRK1A and DYRK1B, respectively. GBM samples were used for Kaplan–Meier survival analysis to evaluate the prognostic power of the expression of DYRK1A and DYRK1B shown

in Fig. 5i. **b**, Distribution of ΔES_{rand} , representing the null model, for ID2 activity (left) and ID2 expression (right). This distribution is used to calculate the P value for enrichment of ΔES . Red dot (or vertical black bar) represents the ΔES using HIF2 α targets. The P value is calculated as ratio of number of times ΔES_{rand} is greater than ΔES (falls in green regions) over the total trials (= 1000).



HIF α degradation

Extended Data Figure 10 | Model for the regulation of HIF α stability by the DYRK1 kinase and ID2 pathway. In cellular contexts that favour HIF α protein instability (normal oxygen levels, but also low ID2 expression and high DYRK1 expression) prolyl hydroxylases (PHD1) is active and positively regulates DYRK1 kinases. Active, tyrosine phosphorylated DYRK1 kinases keep ID2 under functional constraint by phosphorylation of Thr27. The VCB-Cul2 ubiquitin ligase complex efficiently ubiquitylates HIF α (left). With decreasing oxygenation and

HIF α stabilization

PHD1 inactivation but also in the presence of downregulation of DYRK1, elevated expression of ID2, or ID2(T27A) mutation, the un-phosphorylated/un-phosphorylatable pool of ID2 exerts an inhibitory function towards the VCB-Cul2 complex by binding directly VHL and Elongin C proteins and displacing Cul2. This results in HIF α accumulation (right). The transcriptional activation of the *ID2* gene, a HIF α target, by HIF α generates a feed-forward ID2-HIF α loop that amplifies the effects.

Eight per cent leakage of Lyman continuum photons from a compact, star-forming dwarf galaxy

Y. I. Izotov¹, I. Orlitová², D. Schaerer^{3,4}, T. X. Thuan⁵, A. Verhamme³, N. G. Guseva¹ & G. Worseck⁶

One of the key questions in observational cosmology is the identification of the sources responsible for ionization of the Universe after the cosmic ‘Dark Ages’, when the baryonic matter was neutral. The currently identified distant galaxies are insufficient to fully reionize the Universe by redshift $z \approx 6$ (refs 1–3), but low-mass, star-forming galaxies are thought to be responsible for the bulk of the ionizing radiation^{4–6}. As direct observations at high redshift are difficult for a variety of reasons, one solution is to identify local proxies of this galaxy population. Starburst galaxies at low redshifts, however, generally are opaque to Lyman continuum photons^{7–9}. Small escape fractions of about 1 to 3 per cent, insufficient to ionize much surrounding gas, have been detected only in three low-redshift galaxies^{10,11}. Here we report far-ultraviolet observations of the nearby low-mass star-forming galaxy J0925+1403. The galaxy is leaking ionizing radiation with an escape fraction of about 8 per cent. The total number of photons emitted during the starburst phase is sufficient to ionize intergalactic medium material that is about 40 times as massive as the stellar mass of the galaxy.

So-called ‘Green Peas’ (GPs), low-mass compact galaxies with very active star formation^{12–15}, may be promising candidates for sources of escaping ionizing radiation. The GP galaxy J0925+1403 was selected from the Sloan Digital Sky Survey (SDSS) according to the following properties (see Methods for details): (1) a compact structure; (2) the presence of emission lines with high equivalent widths in its SDSS spectrum, suggesting active ongoing star formation and numerous hot O stars producing ionizing Lyman continuum radiation; (3) sufficiently bright in the far-ultraviolet (FUV) with a magnitude of 20.7 mag and redshifted enough ($z = 0.301$) to allow direct Lyman continuum observations with the Cosmic Origins Spectrograph (COS) on board the Hubble Space Telescope (HST); and (4) a high ratio of flux from the [O III] $\lambda = 5,007$ Å line to that from the [O II] $\lambda = 3,727$ Å line, or $O_{32} = [\text{O III}] \lambda 5,007 / [\text{O II}] \lambda 3,727 = 5$ (see Fig. 1), which may indicate the presence of density-bounded H II regions^{14–16}, that is, escaping Lyman continuum radiation.

We first derive some general properties of the galaxy, using the emission-line fluxes measured from the SDSS optical spectrum. After correction for the Milky Way extinction of $A_{V,MW} = 0.084$ mag, we obtain an internal extinction $A_{V,int} = 0.36$ mag, and a low oxygen abundance $12 + \log(\text{O}/\text{H}) = 7.91 \pm 0.03$, or less than 0.2 times the solar value. The details of these determinations are given in Methods section. Here and elsewhere in this Letter, errors are $\pm 1\sigma$.

The same SDSS spectrum is used to fit a spectral energy distribution (SED) to derive the galaxy’s global parameters, including the stellar mass and the age of the present burst of star formation (see Methods). We obtain a starburst age of 2.6 ± 0.2 Myr, a young stellar mass of $(2.4 \pm 0.3) \times 10^8 M_{\odot}$, and a total galaxy stellar mass of $(8.2 \pm 0.7) \times 10^8 M_{\odot}$ (M_{\odot} , solar mass). The star-formation rate is $52.2 M_{\odot} \text{ yr}^{-1}$, as determined from the extinction-corrected H β line flux. With its low mass, low metallicity, low extinction, compact

morphology, and high star-formation rate, J0925+1403 shares many of the properties of high-redshift Lyman- α (Ly α) emitters.

GPs with $O_{32} \geq 5$ have been observed before by HST^{17,18}, but their low redshifts $z < 0.3$ were not optimal for Lyman continuum observations. The HST/COS observations of J0925+1403 were obtained on 28 March 2015 (program GO13744; PI, T.X.T.). The near-ultraviolet acquisition image shows the galaxy to have a very compact structure, with a half-light angular diameter of $\sim 0.2''$, much smaller than the spectroscopic aperture of $2.5''$ (Fig. 2). This angular diameter corresponds to a linear diameter of ~ 1 kpc at the angular diameter distance of 930 Mpc, derived from the redshift $z = 0.301$, adopting the Planck mission cosmological parameters $H_0 = 67.1 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\Lambda} = 0.682$ and $\Omega_{\text{m}} = 0.318$ (ref. 19).

Spectra of J0925+1403 were obtained with two gratings. The low-resolution G140L grating (~ 900 – $2,385$ Å) was used to obtain the spectrum that includes the redshifted Lyman continuum emission, with an exposure time of 5,649 s. The medium-resolution G160M grating ($1,410$ – $1,796$ Å) was used to obtain the spectrum that includes the redshifted Ly α $\lambda 1,216$ Å line, with an exposure time of 2,978 s. The observations with the G160M and G140L gratings were reduced with the standard pipeline and custom software, respectively. The custom software gives more accurate results, as it is specifically designed for faint HST/COS targets (see Methods section).

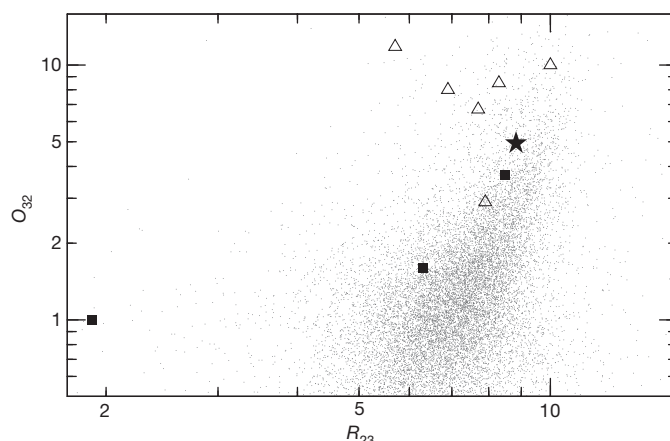


Figure 1 | The O_{32} – R_{23} diagram for star-forming galaxies. The quantity R_{23} is the total flux of the strongest oxygen lines in the optical spectrum relative to H β , and is given by $([\text{O III}] \lambda 4,959 + [\text{O III}] \lambda 5,007 + [\text{O II}] \lambda 3,727) / \text{H}\beta$. This quantity is used for easier comparison with high-redshift Ly α emitting galaxies potentially leaking ionizing radiation¹⁶, shown by open triangles. O_{32} is defined in the main text. At low metallicities, $12 + \log(\text{O}/\text{H}) < 8.3$, R_{23} increases with the metallicity. The location of J0925+1403 and known low-redshift Lyman continuum leaking galaxies^{10,11} are shown by the filled star and filled squares, respectively. SDSS star-forming galaxies²⁷ are represented by small dots.

¹Main Astronomical Observatory, National Academy of Sciences of Ukraine, 27 Zabolotnoho street, Kyiv 03680, Ukraine. ²Astronomical Institute, Czech Academy of Sciences, Boční II 1401, 141 00 Prague, Czech Republic. ³Observatoire de Genève, Université de Genève, 51 Chemin des Maillettes, 1290 Versoix, Switzerland. ⁴CNRS, IRAP, 14 Avenue East Belin, 31400 Toulouse, France. ⁵Astronomy Department, University of Virginia, PO Box 400325, Charlottesville, Virginia 22904, USA. ⁶Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany.

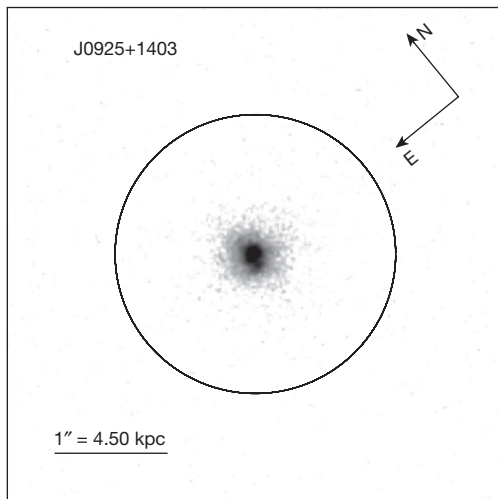


Figure 2 | Near-ultraviolet image of J0925+1403 from the HST. The galaxy with a linear diameter of ~ 2 kpc consists of two compact star-forming regions superimposed on an extended low-surface-brightness component. The spectroscopic aperture with a diameter of $2.5''$ is shown by a circle.

A strong Ly α $\lambda 1,216$ Å emission line is detected in the medium-resolution spectrum. Its profile (Fig. 3) shows two peaks, one on each side of the line centre (dashed vertical line). According to radiative transfer models, the separation between the Ly α line peaks increases with increasing optical depth and thus with increasing neutral hydrogen column density $N(\text{H I})$ (ref. 20). In the case of J0925+1403, the separation of ~ 300 km s $^{-1}$ corresponds to a low column density ($N(\text{H I}) \leq 10^{19}$ cm $^{-2}$), allowing the escape of a considerable fraction of the Ly α emission. Correcting for the Milky Way and galaxy internal reddening, we obtain a Ly α flux density of 8.2×10^{-14} erg s $^{-1}$ cm $^{-2}$. Comparing the extinction-corrected Ly α /H β flux ratio of 16.7 ± 1.0 and case B flux ratio of 23.3 (ref. 21), we find that the Ly α escape fraction is $\sim 70\%$, among the highest known so far for GP galaxies¹⁸, and consistent with a low H I column density.

The short-wavelength part of the J0925+1403 spectrum, obtained with the low-resolution grating G140L, is shown in Fig. 4a by the grey solid line. The modelled ultraviolet SED of the young cluster with the age and extinction parameters obtained before from SED

fitting in the optical range (see Methods) is shown by the black solid line. We adopt the reddening curve from ref. 22, corresponding to $R_{V,MW} = A_V/E_{B-V} = 3.1$ for the Milky Way and to $R_{V,int} = 2.4$ for J0925+1403, except for $\lambda \leq 1,250$ Å, where the reddening curve from ref. 23 is used. The corresponding intrinsic SED is shown in Fig. 4a by the black dash-dotted line. The flux density of the intrinsic Lyman continuum is determined primarily by the extinction-corrected flux density of the H β emission line and the starburst age. The starburst age is derived from the condition that the observed equivalent width of the H β emission line is equal to the modelled value, which depends on the extinction-corrected flux of the continuum near H β and the intrinsic Lyman continuum flux. The intrinsic Lyman continuum is fairly insensitive to the adopted stellar evolution models, stellar atmosphere models and initial mass function (see Methods). It is seen that the reddened SED reproduces the observed spectrum very well for rest-frame wavelengths > 912 Å.

Figure 4b shows a blow-up of the Lyman continuum spectral region. The important feature to note is that the Lyman continuum flux density for $\lambda < 912$ Å is not zero, but positive with a value equal to $(2.35 \pm 0.20) \times 10^{-17}$ erg s $^{-1}$ cm $^{-2}$ Å $^{-1}$, when averaged over the 860–913 Å rest-frame spectral range. It is detected at the 11.6σ level and is indicated by a dotted horizontal line and a filled circle in Fig. 4b. This observed Lyman continuum should be corrected for the Milky Way extinction before the determination of the Lyman continuum escape fraction. The extinction-corrected average Lyman continuum flux density of $(3.43 \pm 0.29) \times 10^{-17}$ erg s $^{-1}$ cm $^{-2}$ Å $^{-1}$ is shown by the thick solid horizontal line in Fig. 4b. Comparing this value to the intrinsic Lyman continuum flux density of 4.4×10^{-16} erg s $^{-1}$ cm $^{-2}$ Å $^{-1}$ beyond the Lyman limit for $\lambda < 912$ Å, we obtain an absolute Lyman continuum escape fraction $f_{esc} = 7.8\% \pm 1.1\%$ in J0925+1403, where the error is determined by the observed Lyman continuum flux density error and uncertainties in the modelled intrinsic Lyman continuum flux density. This value is several times higher than f_{esc} of the other three low-redshift galaxies with known Lyman continuum leakage. J0925+1403 also has the highest [O III]/[O II] flux ratio, the lowest metallicity and the lowest stellar mass. Thus we conclude that compact low-mass star-forming galaxies with high [O III]/[O II] ratios may lose a considerable fraction of their Lyman continuum emission to the intergalactic medium (IGM).

The above determination of f_{esc} holds for ultraviolet-emitting star-forming regions without dust-obscured star formation, which is invisible in the ultraviolet and/or optical ranges. The sky region containing the galaxy J0925+1403 has been observed in the mid-infrared range by the Wide-field Infrared Survey Explorer (WISE). However, data for this galaxy are not present in the AllWISE Source Catalog²⁴. There are also no data in the radio range. Optical, near- and mid-infrared observations of other low-metallicity star-forming galaxies with similar properties suggest that they are relatively transparent²⁵. Negligible dust-obscured star formation is also implied by the observed thermal free-free centimetre radio emission in dwarf galaxies with optical and infrared observations, as it is consistent with the value derived from the flux density of the H β emission line²⁶. Presumably, the same conclusion holds for J0925+1403.

The number of ionizing photons escaping the galaxy is $Q_H = 3.86 \times 10^{53}$ s $^{-1}$ if $f_{esc} = 7.8\%$ (Methods section), corresponding to a total number of ionizing photons of 3.6×10^{67} emitted during a starburst with a 3 Myr duration. This total number of photons is sufficient to ionize the low-density IGM gas with a mass of $\sim 4 \times 10^{10} M_\odot$, or about 40 times higher than the stellar mass of the galaxy, assuming one photon suffices to ionize one hydrogen atom. Here we adopt the luminosity distance of 1,620 Mpc. Finally, we note that our galaxy leaks a large number of ionizing photons, $Q_H/L_{1,500} \approx 10^{25}$ photons s $^{-1}$ /(erg s $^{-1}$ Hz $^{-1}$), per unit ultraviolet luminosity, approximately three times more than (optimistic) assumptions¹ used at high redshift for $f_{esc} = 0.2$, which is primarily due to the young age of the ultraviolet-dominant stellar population in J0925+1403.

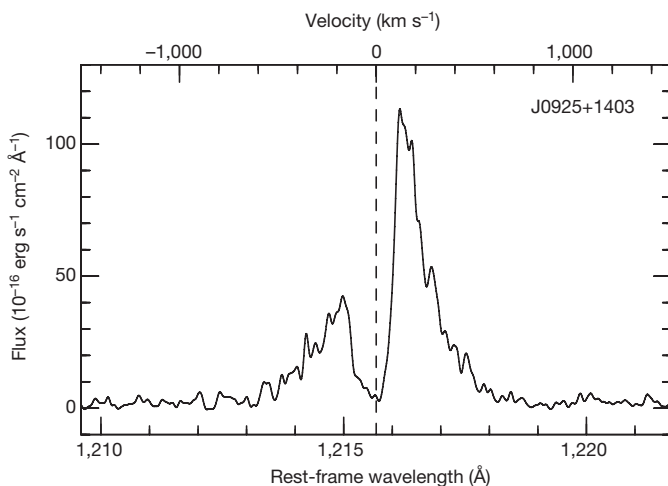


Figure 3 | The double-peaked Ly α emission line in the COS spectrum of J0925+1403. The centre of the emission line is shown by a vertical dashed line. The small separation of the two emission peaks is indicative of a low H I column density according to radiation transfer models for spherical geometry²⁰.

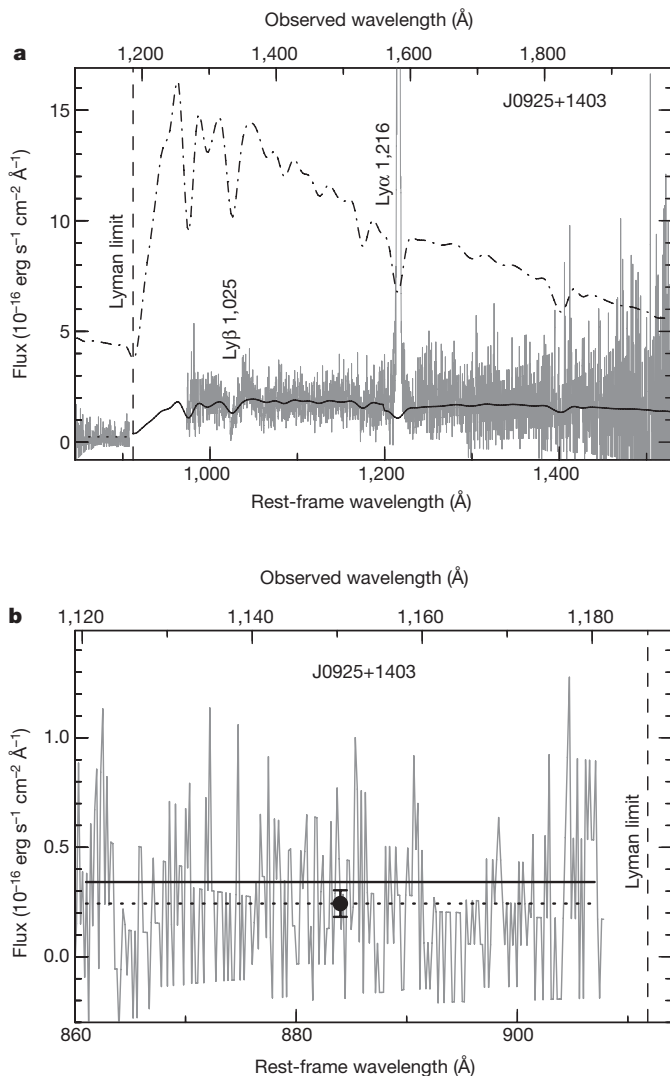


Figure 4 | The COS spectrum of J0925+1403. **a**, On top of the spectrum (grey line) is superposed the modelled SED, reddened by both the Milky Way and internal extinctions (black solid line). The unreddened SED is shown by the dash-dotted line. The short, horizontal dotted line indicates the average observed flux density of the Lyman continuum. **b**, The spectrum in the wavelength range 860–913 Å. The average observed value with $\pm 3\sigma$ error bars is shown by the filled circle and the dotted line. The solid line represents the Lyman continuum level after correction for the Milky Way extinction.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 August; accepted 13 November 2015.

- Robertson, B. E. *et al.* New constraints on cosmic reionisation from the 2012 Hubble Ultra Deep Field Campaign. *Astrophys. J.* **768**, 71 (2013).
- Steidel, C. C., Pettini, M. & Adelberger, K. L. Lyman-continuum emission from galaxies at $z \geq 3.4$. *Astrophys. J.* **546**, 665–671 (2001).
- Iwata, I. *et al.* Detections of Lyman continuum from star-forming galaxies at $z \sim 3$ through Subaru/Suprime-Cam narrow-band imaging. *Astrophys. J.* **692**, 1287–1293 (2009).
- Mitra, S., Ferrara, A. & Choudhury, T. R. The escape fraction of ionising photons from high-redshift galaxies from data-constrained reionisation models. *Mon. Not. R. Astron. Soc.* **428**, L1–L5 (2013).
- Yajima, H., Choi, J.-H. & Nagamine, K. Escape fraction of ionising photons from high-redshift galaxies in cosmological SPH simulations. *Mon. Not. R. Astron. Soc.* **412**, 411–422 (2011).

- Wise, J. H. & Cen, R. Ionising photon escape fractions from high-redshift dwarf galaxies. *Astrophys. J.* **693**, 984–999 (2009).
- Leitherer, C., Ferguson, H. C., Heckman, T. M. & Lowenthal, J. D. The Lyman continuum in starburst galaxies observed with the Hopkins Ultraviolet Telescope. *Astrophys. J.* **454**, L19–L22 (1995).
- Deharveng, J.-M. *et al.* Constraints on the Lyman continuum radiation from galaxies: first results with FUSE on Mrk 54. *Astron. Astrophys.* **375**, 805–813 (2001).
- Grimes, J. P. *et al.* Observations of starburst galaxies with Far-Ultraviolet Spectrographic Explorer: galactic feedback in the Local Universe. *Astrophys. J. Suppl. Ser.* **181**, 272–320 (2009).
- Leitert, E., Bergvall, N., Hayes, M., Linné, S. & Zackrisson, E. Escape of Lyman continuum radiation from local galaxies. Detection of leakage from the young starburst Tol 1247–232. *Astron. Astrophys.* **553**, A106 (2013).
- Borthakur, S., Heckman, T. M., Leitherer, C. & Overzier, R. A. A local clue to the reionisation of the universe. *Science* **346**, 216–219 (2014).
- Cardamone, C. *et al.* Galaxy Zoo Green Peas: discovery of a class of compact extremely star-forming galaxies. *Mon. Not. R. Astron. Soc.* **399**, 1191–1205 (2009).
- Izotov, Y. I., Guseva, N. G. & Thuan, T. X. Green Pea galaxies and cohorts: luminous compact emission-line galaxies in the Sloan Digital Sky Survey. *Astrophys. J.* **728**, 161 (2011).
- Jaskot, A. E. & Oey, M. S. The origin and optical depth of ionising radiation in the “Green Pea” galaxies. *Astrophys. J.* **766**, 91 (2013).
- Stasińska, G., Izotov, Y., Morisset, C. & Guseva, N. Excitation properties of galaxies with the highest [O III]/[O II] ratios. No evidence for massive escape of ionising photons. *Astron. Astrophys.* **576**, A83 (2015).
- Nakajima, K. & Ouchi, M. Ionisation state of inter-stellar medium in galaxies: evolution, SFR– M_{star} – z dependence, and ionising photon escape. *Mon. Not. R. Astron. Soc.* **442**, 900–916 (2014).
- Jaskot, A. E. & Oey, M. S. Linking Ly α and low-ionisation transitions at low optical depth. *Astrophys. J.* **791**, L19 (2014).
- Henry, A., Scarlata, C., Martin, C. S. & Erb, D. Ly α emission from Green Peas: the role of circumgalactic gas density, covering, and kinematics. *Astrophys. J.* **809**, 19 (2015).
- Ade, P. A. *et al.* Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.* **571**, A16 (2014).
- Verhamme, A., Orlitová, I., Schaerer, D. & Hayes, M. Using Lyman- α to detect galaxies that leak Lyman continuum. *Astron. Astrophys.* **578**, A7 (2015).
- Hummer, D. G. & Storey, P. J. Recombination-line intensities for hydrogenic ions – I. Case B calculations for H I and He II. *Mon. Not. R. Astron. Soc.* **224**, 801–820 (1987).
- Cardelli, J. A., Clayton, G. C. & Mathis, J. S. The relationship between infrared, optical, and ultraviolet extinction. *Astrophys. J.* **345**, 245–256 (1989).
- Mathis, J. S. Interstellar dust and extinction. *Annu. Rev. Astron. Astrophys.* **28**, 37–70 (1990).
- Wide-field Infrared Survey Explorer (WISE) and NEOWISE. <http://irsa.ipac.caltech.edu/Missions/wise.html> (2013).
- Izotov, Y. I. & Thuan, T. X. Near-infrared spectroscopy of five blue compact dwarf galaxies: II Zw 40, Mrk 71, Mrk 930, Mrk 996, and SBS 0335–052E. *Astrophys. J.* **734**, 82 (2011).
- Izotov, Y. I., Guseva, N. G., Fricke, K. J., Krügel, E. & Henkel, C. Dust emission in star-forming dwarf galaxies: general properties and the nature of the submm excess. *Astron. Astrophys.* **570**, A97 (2014).
- Izotov, Y. I., Guseva, N. G., Fricke, K. J. & Henkel, C. Multi-wavelength study of 14 000 star-forming galaxies from the Sloan Digital Sky Survey. *Astron. Astrophys.* **561**, A33 (2014).

Acknowledgements This Letter is based on observations made with the NASA/ESA HST, obtained from the data archive at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555. Support for this work was provided by NASA through grant number HST-GO-13744.001-A from the STScI. I.O. acknowledges a grant GACR 14–20666P of the Czech Science Foundation. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. GALEX is a NASA mission managed by the Jet Propulsion Laboratory. This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

Author Contributions All authors participated in the design of the HST observational program. Y.I.I. and N.G.G. selected the galaxy sample. T.X.T. and Y.I.I. led the observations. G.W. reduced the HST data. I.O. did part of the HST data analysis. Y.I.I. and D.S. did the SED modelling and interpretation. A.V. and Y.I.I. did the Ly α interpretation. The bulk of the text was written by Y.I.I. All authors commented on the manuscript at all stages.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.I.I. (izotov@mao.kiev.ua).

METHODS

The sample of compact star-forming galaxies. A sample of compact star-forming galaxies was selected from the spectroscopic data base of the SDSS Data Release 10 (DR10)²⁸ by applying selection criteria as follows: (1) $R_{50} \leq 3''$, where R_{50} is the galaxy's Petrosian radius, within which 50% of the galaxy's flux in the SDSS r band is contained; (2) spiral galaxies were excluded; (3) the emission-line ratio $[\text{O III}]\lambda 4,959/\text{H}\beta$ is ≥ 1 to include only galaxies with high-excitation H II regions, ensuring an accurate determination of their extinctions and chemical compositions; (4) galaxies with AGN (active galactic nuclei) activity were excluded using line ratios as described below. Applying these criteria, 5,182 galaxies were found in the redshift range $0 < z < 1$. Of these, 25 galaxies have $\text{O}_{32} \geq 5$ and are located in the redshift range 0.30–0.35, making direct Lyman continuum observations possible. J0925+1403 with equatorial coordinates RA = 09:25:32.37 and dec. = +14:03:13.06 was selected for the HST/COS ultraviolet observations as one of the brightest objects in the sample of 25 galaxies, with a FUV magnitude of 20.7 mag from GALEX. The high rest-frame equivalent width $\text{EW}_{\text{H}\beta} = 177 \text{ \AA}$ of the $\text{H}\beta$ emission line in the SDSS spectrum of J0925+1403 suggests active ongoing star formation.

The $[\text{O III}]\lambda 5,007/\text{H}\beta$ – $[\text{N II}]\lambda 6,584/\text{H}\alpha$ diagnostic diagram²⁹ of extragalactic objects with emission-line spectra, shown in Extended Data Fig. 1, is useful for discriminating between star-forming galaxies (SFG) and active galactic nuclei (AGN) as ionizing sources. The solid line represents the model line³⁰ that separates the two types of objects. Extended Data Figure 1 clearly shows that the gas in J0925+1403 is ionized by hot stars in star-forming regions.

Custom HST/COS G140L data reduction. J0925+1403 was observed with the HST/COS grating G140L for two orbits at a central wavelength of 1,280 Å in all four focal-plane offset positions to minimize fixed-pattern noise and to patch grid-wire shadows and other detector blemishes. The data were reduced with CALCOS v2.21 and custom software specifically designed for faint HST/COS targets^{31,32}, improving upon previous results on Lyman continuum leakage obtained with the default CALCOS pipeline products¹¹. As our spectra were taken at COS Detector Lifetime Position 3, we did not employ pulse height filtering to preserve source flux in the presence of detector gain sag from previous usage of COS at the nearby COS Lifetime Position 1. We used boxcar extraction in a narrow 25 pixel rectangular window that preserves spectrophotometry for compact sources (Fig. 2) while minimizing the background. The COS background is dominated by detector dark current, which was subtracted in post-processing using scaled dark exposures to account for gain sag in the COS aperture. Specifically, we co-added dark exposures taken in similar orbital conditions within 3 months of the science observations, smoothed the dark current in the COS aperture with a 500-pixel running average to remove Poisson fluctuations, and rescaled the smoothed dark current to the science observations using unilluminated regions of the detector. The resulting estimate of the COS dark current is sufficiently accurate ($\sim 5\%$) over the wavelength range of interest (1,100–1,180 Å) to not affect our analysis. Subexposures were co-added by summing raw counts and the smoothed dark current, accounting for differing pixel exposure times due to detector blemishes before converting to flux via the COS calibration curve. This procedure preserves the Poisson counts of faint sources. Airglow contamination ($\text{N I } \lambda 1,134 \text{ \AA}$ and $\text{O I } \lambda 1,304 \text{ \AA}$) was eliminated by considering only data taken in orbital night in the affected wavelength ranges. We also verified with orbital night data that scattered H I Ly α airglow is negligible in the Lyman continuum of J0925+1403. The sum of diffuse open-shutter backgrounds (earthshine, zodiacal light, Galactic emission³³) is a factor of ~ 30 smaller than the measured flux, leading us to the conclusion that the measured flux is indeed source Lyman continuum flux and not unaccounted background.

Extinction in the optical range. Using the observed decrement of several hydrogen Balmer emission lines, we corrected the line fluxes relative to the $\text{H}\beta$ flux for two effects: (1) reddening adopting the extinction curve from ref. 22 and (2) underlying hydrogen stellar absorption. Both are obtained simultaneously by an iterative procedure³⁴. The quantity derived from the hydrogen Balmer decrement is the extinction coefficient $C_{\text{H}\beta}$, corresponding to the extinction at the $\text{H}\beta$ wavelength $A_{\text{H}\beta} = 2.512 \times C_{\text{H}\beta}$. Using the A_{λ}/A_{V} fits²², we approximate the ratio $C_{\text{H}\beta}/A_{\text{V}}$ by the relation

$$C_{\text{H}\beta}/A_{\text{V}} = 0.6633 - 0.1317R_{\text{V}} + 0.0294R_{\text{V}}^2 - 0.0024R_{\text{V}}^3 \quad (1)$$

with an accuracy better than 0.1% in the range $R_{\text{V}} = A_{\text{V}}/E_{\text{B}-\text{V}} = 2.0$ –4.0. We note that $C_{\text{H}\beta}/A_{\text{V}}$ is weakly dependent on R_{V} , decreasing by $<10\%$ with R_{V} increasing from 2 to 4.

The correction for reddening was done in two steps. First, the observed spectrum, uncorrected for redshift, was corrected for the Milky Way extinction with $A_{\text{V,MW}} = 0.084$ mag and $R_{\text{V,MW}} = 3.1$ (NASA Extragalactic Database), corresponding to $C_{\text{H}\beta,\text{MW}} = 0.039$. Then, the rest-frame spectrum was corrected for the internal extinction of J0925+1403, obtained as $C_{\text{H}\beta,\text{int}} = 0.175$ from the hydrogen Balmer decrement, corresponding to $A_{\text{V,int}} = 0.38$ mag for $R_{\text{V,int}} = 3.1$ and $A_{\text{V,int}} = 0.36$ mag for $R_{\text{V,int}} = 2.4$. The extinction-corrected emission-line fluxes relative to the $\text{H}\beta$

emission line flux, $I(\lambda)/I(\text{H}\beta)$, and the rest-frame equivalent widths, $\text{EW}(\lambda)$, are shown in Extended Data Table 1.

Element abundances. The $[\text{O III}]\lambda 4,363 \text{ \AA}$ emission line is detected in the SDSS spectrum of J0925+1403. This allows for a reliable oxygen abundance determination using the direct T_{e} -method. The temperature $T_{\text{e}}(\text{O III})$ is calculated based on the $[\text{O III}]\lambda 4,363/(\lambda 4,959 + \lambda 5,007)$ line ratio³⁵. We adopt a two-zone photoionized H II region model: a high-ionization zone with temperature $T_{\text{e}}(\text{O III})$, where $[\text{O III}]$ lines originate, and a low-ionization zone with temperature $T_{\text{e}}(\text{O II})$, where $[\text{O II}]$ lines originate. For $T_{\text{e}}(\text{O II})$, we use a model relation between the electron temperatures $T_{\text{e}}(\text{O III})$ and $T_{\text{e}}(\text{O II})$ (ref. 35). Ionic and total abundances of oxygen, nitrogen and neon are derived using expressions for ionic abundances and ionization correction factors (ICF)³⁵. The derived temperatures and element abundances are shown in Extended Data Table 2. We note that the weak $[\text{S II}]\lambda 6,717$, 6,731 emission lines, which are used for the electron number density determination, can not be measured because they are located in the noisy part of the J0925+1403 spectrum. Therefore, we adopted the value $N_{\text{e}} = 100 \text{ cm}^{-3}$, typical of extragalactic H II regions.

Fitting of SED and determination of galaxy global parameters. The luminosity, stellar mass and star-formation rate are important global galaxy characteristics. For the determination of the stellar mass, we follow earlier approach for modelling dwarf galaxies^{13,26,27,36}. The method is based on fitting a series of model SEDs to the observed one and finding the best fit. The fit was performed for the SDSS spectrum over the entire observed spectral range of 3,900–9,200 Å. As the SED is the sum of both stellar and ionized gas emission, its shape depends on the relative contribution of these two components. In J0925+1403, with a rest-frame $\text{H}\beta$ equivalent width $\text{EW}_{\text{H}\beta} = 177 \text{ \AA}$, the ionized gas continuum is strong and should be subtracted before determining the stellar mass.

We carried out a series of Monte Carlo simulations to reproduce the SED of J0925+1403. To derive the stellar SED, we use a grid of instantaneous burst SEDs in a wide range of ages from 0.0 Myr to 15 Gyr, calculated with Starburst99^{37,38}. We adopted Geneva stellar evolution tracks for non-rotating and rotating stars^{39,40} and Padova stellar evolution tracks⁴¹. Various models of stellar atmospheres were used^{42–45}, and we adopt two different stellar initial mass functions^{46,47}. Then the SED with any star-formation history can be obtained by integrating the instantaneous burst SEDs over time with a specified time-varying star-formation rate. The SED of the gaseous continuum was taken into account. It included hydrogen and helium free-bound, free-free and two-photon emission⁴⁸.

The star-formation history is approximated assuming a recent short burst with age $t_{\text{y}} < 10$ Myr, which accounts for the young stellar population, and a prior continuous star formation for the older stars during the time interval between initial time t_i and final time t_f ($t_f < t_i$ and zero age is now). The contribution of each stellar population to the SED was parameterized by the ratio $b = M_{\text{y}}/M_{\text{o}}$, where M_{y} and M_{o} are respectively the masses of the young and old stellar populations. Then the total stellar mass is $M_{\text{s}} = M_{\text{y}} + M_{\text{o}}$.

We calculated 10^5 Monte Carlo models by randomly varying t_{y} , t_i , t_f and b , while other parameters, such as evolutionary tracks, stellar atmosphere models, initial mass function and metallicity are kept fixed. In all cases we used models with metallicities that best match the metallicity of J0925+1403. We also calculate the equivalent width $\text{EW}_{\text{H}\beta}$ for each model. The best solution is required to fulfil two conditions. First, only models in which the modelled equivalent width $\text{EW}_{\text{H}\beta}$ of the $\text{H}\beta$ emission line agrees with the observed value within 5% were selected. Second, the best modelled SED among selected models for each set of fixed parameters was found from χ^2 minimization of the deviation between the modelled and the observed continuum in five wavelength ranges, which are free of the emission lines and residuals of the night-sky lines. We found best solutions for 33 combinations of evolutionary tracks, stellar atmosphere models and initial mass functions to investigate the dependence of the Lyman continuum flux density on the input parameters. All these solutions provide almost equally good fits at rest-frame wavelengths greater than 912 Å and small variations of Lyman continuum. The optical spectrum of J0925+1403 with the overlaid SED of one out of the 33 best models is shown in Extended Data Fig. 2. Here we adopted a Salpeter initial mass function⁴⁶, Padova evolutionary tracks⁴¹ and a combination of stellar atmosphere models^{42,43}.

The modelled SED in the ultraviolet range is shown in Fig. 4a by a dash-dotted line. We note that the amount of the flux decrease at the Lyman break at a fixed starburst age depends on the adopted stellar evolution tracks, stellar atmosphere models, and initial mass function. In particular, the Lyman break in models with non-rotating stars is stronger than in models with rotating stars. However, we find from the SED fitting that the intrinsic flux density $I(\lambda 912)$ of the Lyman continuum varies only by $\sim 10\%$ in the various sets of models. This is primarily due to the fact that to fit the observed $\text{EW}_{\text{H}\beta}$, the starburst age in models with rotating stars should be greater than that in models with non-rotating stars³⁸. For young bursts with an age of 0–3 Myr we obtained the approximate relation $I(\lambda 912) \approx I(\text{H}\beta)/12$, which may be used to estimate the intrinsic Lyman continuum flux density in galaxies

with $f_{\text{esc}} \ll 1$ without modelling. We also note that this relation is not much changed if continuous star formation is adopted for the young stellar population instead of an instantaneous burst. This is due to the dominant contribution of the 0–3 Myr stellar population to both the Lyman continuum and H β fluxes, while the contribution of older stars is small.

To derive global characteristics, the observed fluxes were transformed to luminosities adopting the luminosity distance 1,620 Mpc (ref. 49) and the cosmological parameters $H_0 = 67.1 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.682$, $\Omega_m = 0.318$ (ref. 19). Additionally, for the determination of the H β luminosity $L_{\text{H}\beta}$ and star-formation rate, the H β flux density corrected for aperture effects using the relation $2.512^{r(\text{app})-r}$, where r and $r(\text{app})$ are, respectively, the SDSS r-band total magnitude and the magnitude within the round spectroscopic $3''$ diameter aperture. The aperture correction for J0925+1403 is 1.54. The derived H β luminosity (Extended Data Table 3) corresponds to a number of ionizing photons $Q_{\text{H}} = 4.94 \times 10^{54} \text{ s}^{-1}$, in good agreement with the Q_{H} obtained from the integration of the number of ionizing photons in the modelled intrinsic spectrum (it is only $\sim 8\%$ lower). The corresponding star-formation rate is derived using the standard Kennicutt relation⁵⁰.

Reddening law for J0925+1403. The comparison of the modelled SEDs with the observed photometric and spectroscopic data in the entire ultraviolet and optical range allows to verify which extinction curve is most applicable to J0925+1403. It is known that the extinction curve in the ultraviolet range for the Small Magellanic Cloud with oxygen abundance $12 + \log \text{O}/\text{H} \approx 8.1$ is much steeper than the Milky Way extinction curve, and is characterized by $R_V \approx 2.7$ (refs 51–53). On the other hand, the extinction curve in the optical range is insensitive to variations of R_V . The oxygen abundance of J0925+1403 is lower, ~ 7.9 . Therefore we may expect the extinction curve in this galaxy to be characterized by an even lower R_V .

The observed ultraviolet and optical spectra are shown in Extended Data Fig. 3 by grey lines. Their fluxes are consistent with the SDSS and GALEX photometric fluxes shown by the black symbols. To fit these observational data we reddened the intrinsic modelled SED, keeping the same Milky Way extinction $C_{\text{H}\beta, \text{MW}} = 0.039$, $R_{V, \text{MW}} = 3.1$, and the internal extinction $C_{\text{H}\beta, \text{int}} = 0.175$, but varying $R_{V, \text{int}} = 3.1$ (dotted line), 2.7 (dashed line) and 2.4 (solid line). The modelled SEDs were calculated with Padova stellar evolution models⁴¹, a combination of stellar atmosphere models^{42,43} and a Salpeter initial mass function⁴⁶. It is seen that the SEDs reddened with $R_{V, \text{int}} = 3.1$ and 2.7 do not fit the observed fluxes in the ultraviolet range. A higher extinction coefficient $C_{\text{H}\beta, \text{int}}$ would be needed in the ultraviolet. This is difficult to understand in the framework of a model with a uniform dust distribution, characterized by a single value of the extinction. On the other hand, a SED reddened by the same extinction $C_{\text{H}\beta, \text{int}} = 0.175$, but with $R_{V, \text{int}} = 2.4$, nicely reproduces the observed fluxes. If we use different stellar atmosphere models for the SED modelling, we find somewhat larger $R_{V, \text{int}}$ in some cases, but not exceeding 2.7.

Relative Lyman continuum escape fraction. To compare with other studies, we also estimate the Lyman continuum escape fraction from the often-used equation¹⁰.

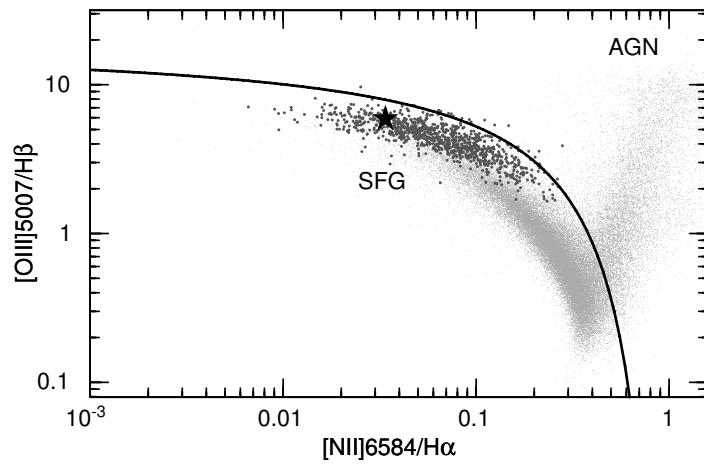
$$f_{\text{esc}} = f_{\text{esc, rel}} \times 10^{-0.4 \times A_{1,500}} = \frac{(f_{1,500} / f_{900})_{\text{int}}}{(f_{1,500} / f_{900})_{\text{obs}}} \times 10^{-0.4 \times A_{1,500}} \quad (2)$$

where f_{900} and $f_{1,500}$ are the flux densities at rest-frame wavelengths 900 Å and 1,500 Å, respectively, the subscripts ‘int’ and ‘obs’ denote mean intrinsic and observed flux densities, $(f_{1,500} / f_{900})_{\text{int}} = 1.36$ from our SED fits, $(f_{1,500} / f_{900})_{\text{obs}} = 5.79$, and $A_{1,500} = 1.58$ mag is the internal extinction at 1,500 Å (Fig. 4a). From this we obtain the relative escape fraction $f_{\text{esc, rel}} = 23.6\%$, and the absolute $f_{\text{esc}} = 5.5\%$. The latter value is lower than our above determination for two reasons. First, equation (2) does not take into account the fact that the expected extinction at 900 Å is higher than at 1,500 Å. And second, the observed flux density f_{900} must be corrected for foreground extinction from the Milky Way. Therefore, f_{esc} obtained with equation (2) is underestimated.

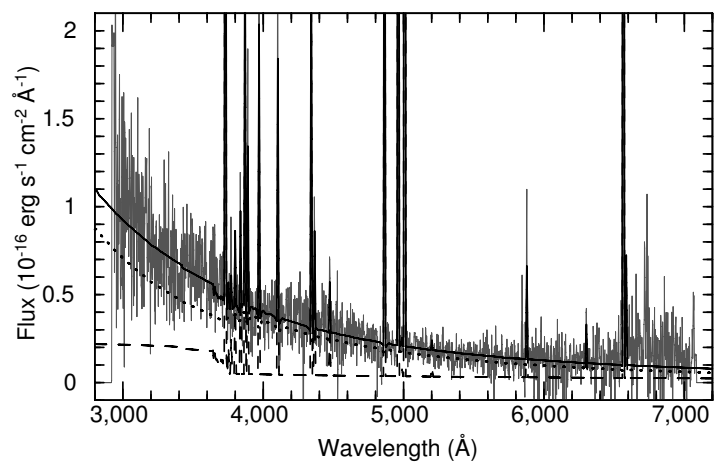
Code availability. We have opted not to make the codes for the custom HST/COS data reduction³¹ and the galaxy SED fitting²⁸ available because they are not yet adapted for public use. The Starburst99 code used to generate spectral

energy distribution of single stellar populations is available at <http://www.stsci.edu/science/starburst99/docs/parameters.html>.

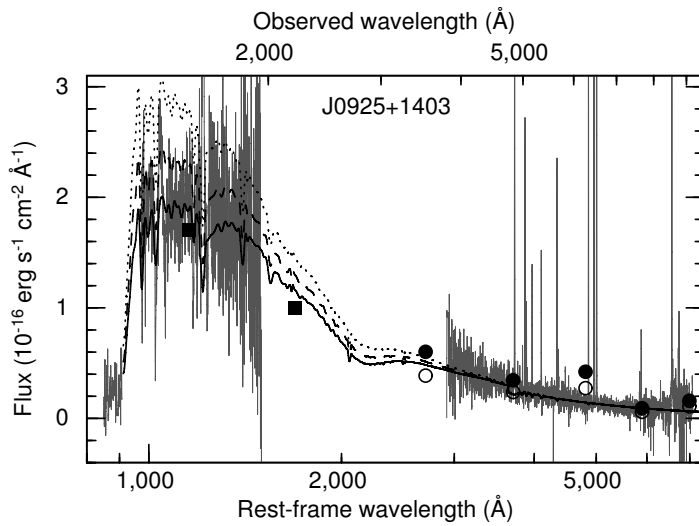
28. Izotov, Y. I., Guseva, N. G., Fricke, K. J. & Henkel, C. On the universality of luminosity-metallicity and mass-metallicity relations for compact star-forming galaxies at redshifts $0 < z < 3$. *Mon. Not. R. Astron. Soc.* **451**, 2251–2262 (2015).
29. Baldwin, J. A., Phillips, M. M. & Terlevich, R. Classification parameters for the emission-line spectra of extragalactic objects. *Publ. Astron. Soc. Pacif.* **93**, 5–19 (1981).
30. Kauffmann, G. et al. Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey. *Mon. Not. R. Astron. Soc.* **341**, 33–53 (2003).
31. Worseck, G. et al. The end of helium reionization at $z \approx 2.7$ inferred from cosmic variance in HST/COS He II Ly α absorption spectra. *Astrophys. J.* **733**, L24 (2011).
32. Syphers, D. et al. HST/COS observations of thirteen new He II quasars. *Astron. J.* **143**, 100 (2012).
33. Murthy, J. GALEX diffuse observations of the sky: the data. *Astrophys. J. Suppl. Ser.* **213**, 32 (2014).
34. Izotov, Y. I., Thuan, T. X. & Lipovetsky, V. A. The primordial helium abundance from a new sample of metal-deficient blue compact galaxies. *Astrophys. J.* **435**, 647–667 (1994).
35. Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G. & Thuan, T. X. The chemical composition of metal-poor emission-line galaxies in the Data Release 3 of the Sloan Digital Sky Survey. *Astron. Astrophys.* **448**, 955–970 (2006).
36. Guseva, N. G., Izotov, Y. I. & Thuan, T. X. Balmer and Paschen jump temperature determinations in low-metallicity emission-line galaxies. *Astrophys. J.* **644**, 890–906 (2006).
37. Leitherer, C. et al. Starburst99: synthesis models for galaxies with active star formation. *Astrophys. J. Suppl. Ser.* **123**, 3–40 (1999).
38. Leitherer, C. et al. The effects of stellar rotation. II. A comprehensive set of Starburst99 models. *Astrophys. J. Suppl. Ser.* **212**, 14 (2014).
39. Meynet, G., Maeder, A., Schaller, G., Schaerer, D. & Charbonnel, C. Grids of massive stars with high mass loss rates. V. From 12 to $120 M_\odot$ at $Z = 0.001$, 0.004, 0.008, 0.020 and 0.040. *Astron. Astrophys. Suppl. Ser.* **103**, 97–105 (1994).
40. Ekström, S. et al. Grids of stellar models with rotation. I. Models from 0.8 to $120 M_\odot$ at solar metallicity ($Z = 0.014$). *Astron. Astrophys.* **537**, A146 (2012).
41. Girardi, L., Bressan, A., Bertelli, G. & Chiosi, C. Evolutionary tracks and isochrones for low- and intermediate-mass stars: from 0.15 to $7 M_\odot$, and from $Z = 0.0004$ to 0.03. *Astron. Astrophys. Suppl. Ser.* **141**, 371–383 (2000).
42. Lejeune, T., Buser, R. & Cuisinier, F. Standard stellar library for evolutionary synthesis. I. Calibration of theoretical spectra. *Astron. Astrophys. Suppl. Ser.* **125**, 229–246 (1997).
43. Schmutz, W., Leitherer, C. & Gruenwald, R. Theoretical continuum energy distributions for Wolf-Rayet stars. *Publ. Astron. Soc. Pacif.* **104**, 1164–1172 (1992).
44. Hillier, D. J. & Miller, D. L. The treatment of non-LTE line blanketing in spherically expanding outflows. *Astrophys. J.* **496**, 407–427 (1998).
45. Pauldrach, A. W. A. et al. Realistic Models for Expanding Atmospheres. In *ASP Conf. Ser. Vol. 131, Properties of Hot, Luminous Stars* (ed. Howarth, I. D.) 258–277 (Astronomical Society of the Pacific, 1998).
46. Salpeter, E. E. The luminosity function and stellar evolution. *Astrophys. J.* **121**, 161–167 (1955).
47. Kroupa, P. On the variation of the initial mass function. *Mon. Not. R. Astron. Soc.* **322**, 231–246 (2001).
48. Aller, L. H. *Physics of Thermal Gaseous Nebulae* (Astrophysics and Space Science Library Vol. 112, Reidel, 1984).
49. Wright, E. L. A cosmology calculator for the World Wide Web. *Publ. Astron. Soc. Pacif.* **118**, 1711–1715 (2006).
50. Kennicutt, R. C. Jr. Star formation in galaxies along the Hubble sequence. *Annu. Rev. Astron. Astrophys.* **36**, 189–231 (1998).
51. Bouchet, P., Lequeux, J., Maurice, E., Prévot, L. & Prévot-Burnichon, M. L. The visible and infrared extinction law and the gas-to-dust ratio in the Small Magellanic Cloud. *Astron. Astrophys.* **149**, 330–336 (1985).
52. Gordon, K. D. & Clayton, G. C. Starburst-like dust extinction in the Small Magellanic Cloud. *Astrophys. J.* **500**, 816–824 (1998).
53. Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U. & Wolff, M. J. A quantitative comparison of the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way ultraviolet to near-infrared extinction curves. *Astrophys. J.* **594**, 279–293 (2003).



Extended Data Figure 1 | The diagnostic diagram for narrow emission lines. The foundations of this diagram are given in ref. 29. The galaxy J0925+1403 is shown by a large filled star, and the Luminous Compact Galaxies¹³ by small dark-grey circles. Also plotted are the 100,000 emission-line galaxies from SDSS DR7 (cloud of light-grey dots). The solid line³⁰ separates star-forming galaxies (SFG) from active galactic nuclei (AGN).



Extended Data Figure 2 | SED fitting of the optical spectrum of J0925+1403. The rest-frame extinction-corrected spectrum is shown by a grey line. The stellar, ionized gas, and total modelled SEDs are shown by black dotted, dashed and solid lines, respectively.



Extended Data Figure 3 | A comparison of the observed ultraviolet and optical spectrum with the modelled SED. The observed spectrum is shown by a grey line. The total GALEX and SDSS photometric fluxes are represented by filled squares and filled circles, respectively, while the SDSS photometric fluxes within a round spectroscopic aperture of $3''$ diameter are shown by open circles. Modelled SEDs, which are reddened by the Milky Way with $R_{V,MW} = 3.1$ and internal extinction with different values of $R_{V,int}$, are shown by black lines. Dotted, dashed and solid lines correspond to $R_{V,int} = 3.1$, 2.7 , and 2.4 , respectively.

Extended Data Table 1 | Emission-line fluxes and equivalent widths in the optical spectrum

Line	Wavelength (Å)	$100 \times I(\lambda)/I(\text{H}\beta)^{\dagger}$	$\text{EW}(\lambda)^{\ddagger}$ (Å)
[OII]	3727	125.5 ± 4.6	77
H9	3835	9.1 ± 2.0	7
[NeIII]	3868	47.5 ± 1.9	33
HeI+H8	3889	17.1 ± 2.3	12
[NeIII]+H7	3968	32.0 ± 2.6	22
H δ	4101	28.7 ± 2.3	22
H γ	4340	43.7 ± 2.5	36
[OIII]	4363	11.7 ± 0.6	10
HeI	4471	6.1 ± 0.5	6
H β	4861	100.0 ± 3.6	177
[OIII]	4959	199.9 ± 6.7	306
[OIII]	5007	$608.1 \pm 20.$	1174
HeI	5876	11.3 ± 0.7	11
[OI]	6300	4.6 ± 0.5	11
H α	6563	280.2 ± 9.9	732
[NII]	6584	13.2 ± 0.8	26

[†]Extinction-corrected flux relative to the extinction-corrected flux $I(\text{H}\beta) = 4.92 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$ of the H β emission line, multiplied by 100.

[‡]Rest-frame equivalent width.

Extended Data Table 2 | Physical conditions and chemical composition

Parameter	Value
$T_e(\text{OIII}), \text{K}$	15010 ± 410
$T_e(\text{OII}), \text{K}$	14010 ± 360
$N_e(\text{OII}), \text{K}$	100^\dagger
$\text{O}^+/\text{H}^+ \times 10^5$	1.42 ± 0.11
$\text{O}^{2+}/\text{H}^+ \times 10^5$	6.65 ± 0.50
$\text{O}/\text{H} \times 10^5$	8.06 ± 0.52
$12 + \log \text{O}/\text{H}$	7.91 ± 0.03
$\text{N}^+/\text{H}^+ \times 10^6$	1.12 ± 0.07
$\text{ICF}(\text{N})^\ddagger$	5.42
$\text{N}/\text{H} \times 10^6$	6.05 ± 0.45
$\log \text{N}/\text{O}$	-1.12 ± 0.04
$\text{Ne}^{2+}/\text{H}^+ \times 10^5$	1.27 ± 0.11
$\text{ICF}(\text{Ne})^\ddagger$	1.08
$\text{Ne}/\text{H} \times 10^5$	1.37 ± 0.13
$\log \text{Ne}/\text{O}$	-0.77 ± 0.05

[†]Assumed value.[‡]Ionization correction factor.

Extended Data Table 3 | Global characteristics of J0925+1403

Parameter	Value
$I_{\text{H}\beta}^{\dagger}$	49.2 ± 1.3
Redshift	0.301323
Luminosity distance [‡]	1620
$L_{\text{H}\beta}^{\dagger\dagger}$	$(2.32 \pm 0.04) \times 10^{42}$
SFR ^{‡‡}	52.2
Q_{H}^*	4.94×10^{54}
$Q_{\text{H}}(\text{esc})^*$	3.86×10^{53}
$t(\text{burst})^{**}$	2.6 ± 0.2
M_{y}/M_{\odot}	$(2.4 \pm 0.3) \times 10^8$
M_{\star}/M_{\odot}	$(8.2 \pm 0.7) \times 10^8$

[†]Extinction-corrected flux density in units of $10^{-16} \text{ erg s}^{-1} \text{ cm}^{-2}$.

[‡]In units of Mpc.

^{††}Extinction- and aperture-corrected luminosity in units of erg s^{-1} .

^{‡‡}Star-formation rate in $M_{\odot} \text{ yr}^{-1}$ derived from the $\text{H}\beta$ luminosity⁵⁰.

* Q_{H} and $Q_{\text{H}}(\text{esc})$ are the numbers of Lyman continuum photons (in units of s^{-1}) emitted by massive stars and escaped from the H II region, respectively.

**Burst age in Myr.

Weakened magnetic braking as the origin of anomalously rapid rotation in old field stars

Jennifer L. van Saders^{1,2,3}, Tugdual Ceillier⁴, Travis S. Metcalfe⁵, Victor Silva Aguirre⁶, Marc H. Pinsonneault^{3,7}, Rafael A. García^{3,4}, Savita Mathur^{3,5} & Guy R. Davies^{6,8}

A knowledge of stellar ages is crucial for our understanding of many astrophysical phenomena, and yet ages can be difficult to determine. As they become older, stars lose mass and angular momentum, resulting in an observed slowdown in surface rotation¹. The technique of ‘gyrochronology’ uses the rotation period of a star to calculate its age^{2,3}. However, stars of known age must be used for calibration, and, until recently, the approach was untested for old stars (older than 1 gigayear, Gyr). Rotation periods are now known for stars in an open cluster of intermediate age⁴ (NGC 6819; 2.5 Gyr old), and for old field stars whose ages have been determined with asteroseismology^{5,6}. The data for the cluster agree with previous period–age relations⁴, but these relations fail to describe the asteroseismic sample⁷. Here we report stellar evolutionary modelling^{5,6,8–10}, and confirm the presence of unexpectedly rapid rotation in stars that are more evolved than the Sun. We demonstrate that models that incorporate dramatically weakened magnetic braking for old stars can—unlike existing models—reproduce both the asteroseismic and the cluster data. Our findings might suggest a fundamental change in the nature of ageing stellar dynamos, with the Sun being close to the critical transition to much weaker magnetized winds. This weakened braking limits the diagnostic power of gyrochronology for those stars that are more than halfway through their main-sequence lifetimes.

There are two approaches to the calibration and testing of gyrochronology. The first is a purely empirical approach, which uses a sample of stars with independently measured ages and rotation periods to construct period–age relationships. These relationships are generally simple power laws that take into account age, period, and some mass-dependent quantity; they have seen wide usage^{1,2,4,5,7}. The second, model-based approach uses stellar models and a prescription for magnetic braking to account for the functional dependence of the rotation period on all relevant stellar quantities, but relies on calibrators to determine the magnitude of the angular momentum loss. For this reason, the model-based approach is well suited to calibrating samples that cover parameter space only sparsely; it also provides a method for attaching physical meaning to observed braking behaviour.

Magnetic-braking prescriptions are typically scaled from the solar case; for example, the Skumanich relation¹ yields angular momentum loss of the form $dJ/dt \propto \omega^3$, where t is time, J is angular momentum, and ω is the angular rotation velocity¹¹. These relations often use the dimensionless Rossby number—defined as the ratio of the rotation period to the convective overturn timescale, $Ro = P/\tau_{\text{co}}$ —to characterize departures from this simple power law. Rossby-number thresholds and scalings are routinely invoked to parameterize the magnetic-field strength^{12,13}; the dependence of the spin-down on stellar mass and

composition^{2,14}; the observed saturation of magnetic braking in rapid rotators; and the sharp transition from slow to rapid rotation that occurs in hot stars (of greater than 6,250 K) because of their thinning convective envelopes¹⁴. Under traditional prescriptions, stars undergo braking throughout their main-sequence lifetimes, regardless of rotation rate. Observations of stellar clusters of young and intermediate ages have indicated that such treatments are reasonable^{4,15}. However, there is a dearth of old stars with which to test such relationships, owing to the long-period, low-amplitude signatures of rotation in such stars, and to the challenge of age measurements in field stars. Data from the Kepler telescope provide a first test of these prescriptions in stars that are older than the Sun.

The high-precision, long-baseline light curves from Kepler make such investigations possible. The rotation of a star manifests itself in Kepler data as a periodic modulation in the intensity, as dark starspots rotate into and out of view. Intensity variations due to stellar oscillations are likewise present in the light curve, on shorter timescales. Low-degree modes of oscillation probe the conditions of the deep stellar interior and internal structure of the star, providing ages that are precise to better than 10% in stars for which many oscillation modes are detected at high signal-to-noise ratios¹⁶.

The first efforts to calibrate the gyrochronology relations using Kepler seismic targets uncovered tension between the cluster and seismic samples⁷. Although the form of the mass–period–age relation used in this study⁷ was similar to those in previous studies^{2,4}, the range of ages and more sophisticated treatment of observational uncertainties made it possible to determine that the sample did not obey a single power-law period–age relation. However, even this approach has limitations: it does not account for metallicity or for changes in the stellar moment of inertia, and it relied on a sample for which detailed seismic modelling and spectroscopic data were lacking for some stars, biasing the seismic ages.

To address the limitations of previous work and to take full advantage of precisely determined stellar parameters, we utilize a subset of 21 Kepler stars—selected to have detailed asteroseismic modelling and high-precision ages, measured rotation periods, and measured metallicities^{5,6,8–10}—and couple these observations to stellar evolutionary models. Sample selection, details of the modelling to derive asteroseismic ages, and extraction of rotation periods are described in the Methods. Figure 1 shows the surface—in terms of period, age and effective temperature (T_{eff} , a proxy for mass)—upon which the stars are expected to lie^{2,4}. Actual cluster and seismic data are overplotted; while the clusters and young asteroseismic targets lie close to the plane, the intermediate-age and old asteroseismic stars are strikingly discrepant and nearly all lie below the surface, owing to the fact that they are rotating more rapidly than expected. When we account

¹Carnegie Observatories, 813 Santa Barbara Street, Pasadena, California 91101, USA. ²Department of Astrophysical Sciences, Princeton University, Princeton, New Jersey 08544, USA.

³Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106-4030, USA. ⁴Laboratoire AIM, Alternative Energies and Atomic Energy Commission (CEA)/Physical Sciences Division (DSM)—Centre National d’Études Spatiales (CNRS)—Université Paris Diderot—Institute of Research into the Fundamental Laws of the Universe (IRFU)/Service d’Astrophysique (SAp), Centre de Saclay, 91191 Gif-sur-Yvette Cedex, France. ⁵Space Science Institute, 4750 Walnut Street Suite 205, Boulder, Colorado 80301, USA. ⁶Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus, Denmark. ⁷Ohio State University, Department of Astronomy, 140 West 18th Avenue, Columbus, Ohio 43210, USA.

⁸School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK.

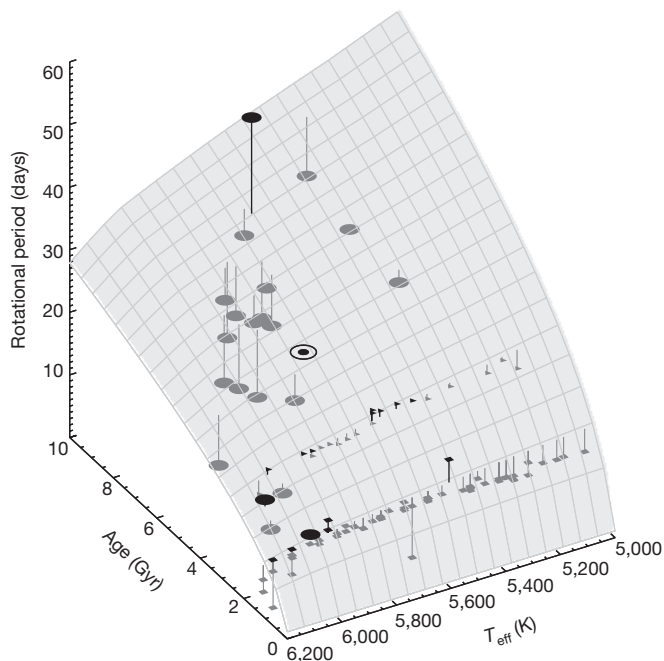


Figure 1 | The period–age plane as predicted by gyrochronology, compared with observed periods. The empirical gyrochronology relation^{2,4} is shown as a plane. Data from open stellar clusters are shown as small squares (NGC 6811 cluster; 1 Gyr) and triangles (NGC 6819 cluster; 2.5 Gyr). Large circles represent the seismic sample of 21 stars that are detected in the Kepler data; this sample falls systematically below the plane. The solar symbol (⊙) marks the Sun, which falls on the plane by design. The effective temperature, T_{eff} , is a proxy for mass.

for uncertainties in the ages, masses, and compositions (see Methods) and predict the rotation periods that we should have observed given existing period–age relations^{2,14} (P_{expected}), we find that the systematic offset persists; stars of roughly solar age and older are rotating more rapidly than predicted, regardless of the chosen period–age relation. Figure 2 highlights the systematic offset by plotting the ratios of the expected to observed periods for each star in the sample, where the expected periods are calculated using stellar models with a braking law calibrated on the Sun and on open clusters¹⁴ (a similar plot is provided in Extended Data Fig. 3 for the empirical relation²). The theoretical models¹⁴ fit the data with a χ^2 value of 54.9, whereas the empirical relation² yields a χ^2 of 155.6. In both cases, the systematic offset towards short rotation periods is an indication that the models predict more angular momentum loss than actually occurs.

We therefore conclude that magnetic braking is weaker in these intermediate-age and old stars. We extend our model by postulating that, in addition to the Rossby scaling already present in the theoretical models¹⁴, effective loss of angular momentum ceases above a critical Rossby threshold¹². We modify the prescription for angular momentum loss¹⁴ to conserve angular momentum above a specified Ro_{crit} . Graphs showing the effects of varying Ro_{crit} values on the models are provided in Fig. 3. The inclusion of the threshold has the desired effect: it reproduces the existing gyrochronology relations and cluster data at young ages, when Ro is smaller because of more rapid rotation, but allows stars to maintain unusually rapid rotation periods at late times. Furthermore, it reproduces the trend in mass that is apparent in Figs 2 and 3 (and the trend in the zero-age main-sequence (ZAMS) T_{eff} , which selects stars with similar rotational histories; we perform all fits using the seismic mass, but use ZAMS T_{eff} for display to simplify the figures). Hotter, more massive stars reach the critical Rossby threshold at younger ages, and we therefore see discrepancies between the fiducial gyrochronology relationships and the observations at

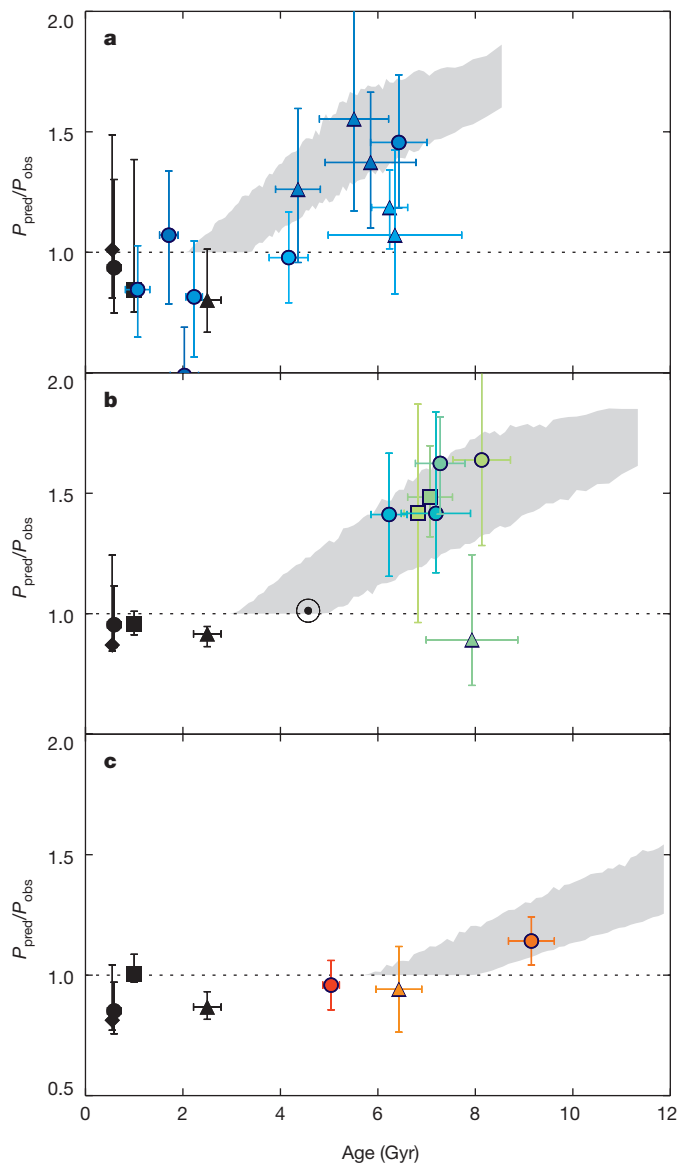


Figure 2 | Ratios of the predicted rotation period¹⁴ to the observed period. Predicted rotation periods are derived from existing period–age relations; observed periods are as detected by the Kepler telescope. These ratios are plotted against stellar age. Stars are divided according to decreasing ZAMS T_{eff} : **a**, 5,900–6,200 K; **b**, 5,600–5,900 K; **c**, 5,100–5,400 K. Period ratios for open clusters are shown as black symbols, as follows: diamonds, M37; circles, Praesepe; squares, NGC 6811; triangles, NGC 6819. The Sun (⊙) is also marked. Coloured circles represent seismic targets; coloured triangles represent known planet hosts; coloured squares represent the binary stars 16 Cygni A and B. All errors are shown to 1σ . Stars are coloured according to ZAMS T_{eff} , with blue representing the hottest stars and red the coolest stars. Shaded regions represent the period ratios permitted in each T_{eff} bin for a model in which $\text{Ro}_{\text{crit}} = 2.16$.

earlier times as ZAMS T_{eff} increases. The best-fit value for the Rossby threshold, given our sample, is $\text{Ro}_{\text{crit}} = 2.16 \pm 0.09$ ($\chi^2 = 13.3$) for the modified models. The shaded grey regions in Figs 2 and 3 denote the full range of period ratios ($P_{\text{Ro}_{\text{crit}}}/P_{\text{fiducial}}$), and the period–age combinations allowed for a model with $\text{Ro}_{\text{crit}} = 2.16$, given the ranges of ZAMS T_{eff} that are represented in each panel. These regions encompass all combinations of mass (0.4–2.0 solar masses) and metallicity ($-0.4 < [Z/H] < +0.4$) that together produce a star within the appropriate ZAMS T_{eff} range for each panel of Figs 2 and 3, on both the main-sequence and the subgiant branch.

We emphasize that our result—that old stars are rotating anomalously rapidly—persists regardless of the choice of period–age

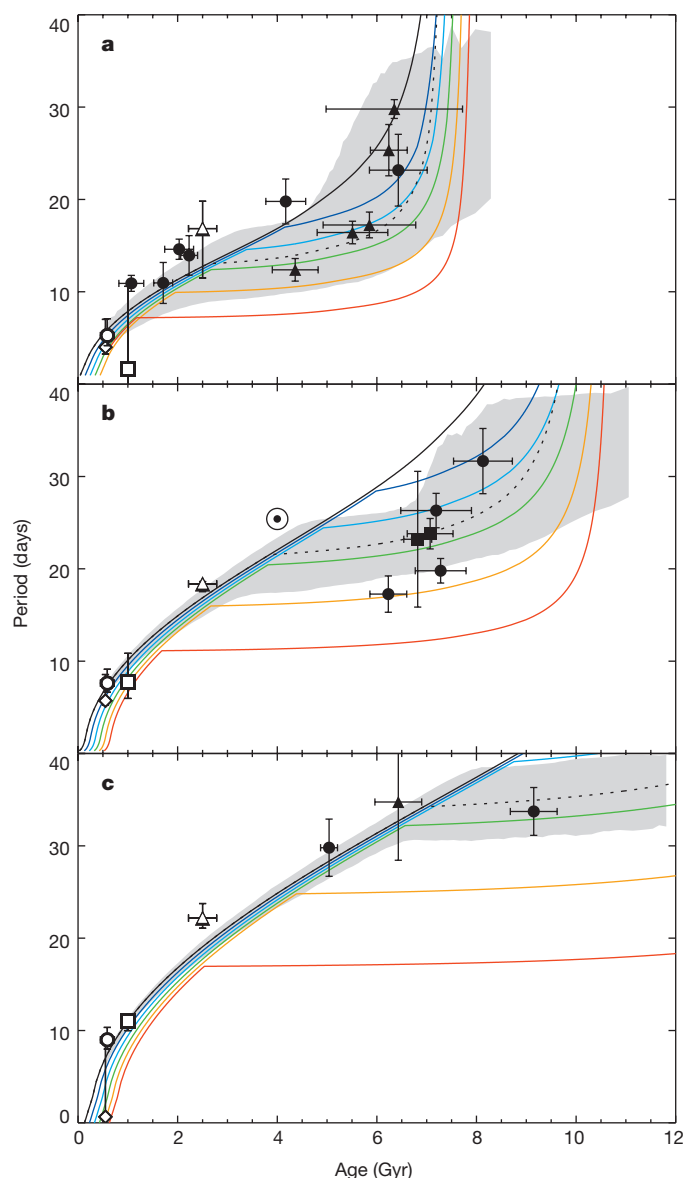


Figure 3 | The effects of a Ro_{crit} threshold on rotational evolution. Panels are divided according to decreasing ZAMS T_{eff} : **a**, 5,900–6,200 K; **b**, 5,600–5,900 K; **c**, 5,100–5,400 K (as in Fig. 2). Black symbols represent open stellar clusters, as follows: diamonds, M37; circles, Praesepe; squares, NGC 6811; triangles, NGC 6819. The Sun (\odot) is also marked. Model curves are shown for solar metallicity and ZAMS T_{eff} 6,050 K (**a**), 5,750 K (**b**), and 5,250 K (**c**). Curves are colour-coded by Ro_{crit} : black, no Ro_{crit} cut; dark blue, $Ro_{crit} = 1.0$; light blue, $Ro_{crit} = 1.5$; green, $Ro_{crit} = 2.0$; orange, $Ro_{crit} = 2.5$; red, $Ro_{crit} = 3.0$; dashed black, $Ro_{crit} = 2.16$. Successive curves are offset by +0.1 Gyrs to improve readability. Seismic (cluster) targets are overplotted in solid (open) symbols with 1σ errors. Shaded regions represent $Ro_{crit} = 2.16$ models for each T_{eff} range.

relationship, asteroseismic modelling pipeline, or model uncertainties from the literature (see Methods). The period-detection algorithms¹⁷ and seismic ages have been well tested⁸. The tight rotational sequences observed in intermediate-age open clusters⁴ suggest that we are not simply detecting the rapidly rotating tail of a population with a wide distribution of rotation rates, and it is unlikely that our 21 stars with detected rotation rates are atypical (see Methods for further discussion).

Our model represents the limiting case in which the braking is so ineffective that the star ceases to shed angular momentum. If we instead allow the exponent, α , of the period–age relation $P \propto t^{1/\alpha}$ to vary, while fixing Ro_{crit} to the solar Ro value of 2.16, we do not obtain a comparable

fit in the old stars until α is greater than ~ 20 , suggesting that the braking is indeed drastically reduced. However, we do observe spot modulation in these stars, which implies at least small-scale magnetic activity. The starspot properties may or may not directly reflect changes in the large-scale magnetic field that governs spin-down. A change in field geometry from a simple dipole to higher-order fields could produce weakened braking^{18,19}, as could a change in the distribution of spots on the stellar surface²⁰. It could also be the case that the large-scale field strength undergoes a transition at high Rossby numbers¹². Abrupt changes in the efficiency of angular momentum loss have been proposed in order to explain the rotational distributions in young clusters²¹, and there is evidence for a Rossby-number-governed shift in field morphologies in low-mass M dwarfs²². Observations of detailed magnetic-field morphologies and corresponding simulations are lacking for stars at higher Rossby numbers than the Sun, and both are critical to understanding the source of the observed anomalous rotation.

Regardless of the mechanism that governs the spin-down, the observation that existing rotation–age relationships do not predict the observed rotation rates has immediate implications for gyrochronology. The rotation periods of the middle-aged stars that have passed this Rossby threshold represent only lower limits on the age. The empirical calibrations must be modified, and the weakened relationship between period and age will result in substantially more uncertain rotation-based ages for stars in the latter halves of their lives. The presence of such a Rossby threshold defines boundaries in mass–age space, past which gyrochronology is incapable of delivering precise ages.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 June; accepted 15 October 2015.

Published online 4 January; corrected online 13 January 2016 (see full-text HTML version for details).

- Skumanich, A. Time scales for CA II emission decay, rotational braking, and lithium depletion. *Astrophys. J.* **171**, 565–567 (1972).
- Barnes, S. A. A simple nonlinear model for the rotation of main-sequence cool stars. I. Introduction, implications for gyrochronology, and color-period diagrams. *Astrophys. J.* **722**, 222–234 (2010).
- Epstein, C. R. & Pinsonneault, M. H. How good a clock is rotation? The stellar rotation-mass-age relationship for old field stars. *Astrophys. J.* **780**, 159–183 (2014).
- Meibom, S. *et al.* A spin-down clock for cool stars from observations of the 2.5-billion-year-old cluster. *Nature* **517**, 589–591 (2015).
- García, R. A. Rotation and magnetism of *Kepler* pulsating solar-like stars. Towards asteroseismically calibrated age-rotation relations. *Astron. Astrophys.* **572**, A34–A49 (2014).
- Ceillier, T. *et al.* Rotation periods and seismic ages of KOIs—comparison with stars without detected planets from *Kepler* observations. *Mon. Not. R. Astron. Soc.* **456**, 119–125 (2016).
- Angus, R. & Aigrain, S., Foreman-Mackey, D. & McQuillan, A. Calibrating gyrochronology using *Kepler* asteroseismic targets. *Mon. Not. R. Astron. Soc.* **450**, 1787–1798 (2015).
- Mathur, S. *et al.* A uniform asteroseismic analysis of 22 solar-type stars observed by *Kepler*. *Astrophys. J.* **749**, 152 (2012).
- Metcalfe, T. S. *et al.* Properties of 42 solar-type *Kepler* targets from the asteroseismic modeling portal. *Astrophys. J.* **214** (Suppl.), 27–40 (2014).
- Silva Aguirre, V. Ages and fundamental properties of *Kepler* exoplanet host stars from asteroseismology. *Mon. Not. R. Astron. Soc.* **452**, 2127–2148 (2015).
- Kawaler, S. D. Angular momentum loss in low-mass stars. *Astrophys. J.* **333**, 236–247 (1988).
- Durney, B. R. & Latour, J. On the angular momentum loss of late-type stars. *Geophys. Astrophys. Fluid Dyn.* **9**, 241–255 (1977).
- Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K. & Vaughan, A. H. Rotation, convection, and magnetic activity in lower main-sequence stars. *Astrophys. J.* **279**, 763–777 (1984).
- van Saders, J. L. & Pinsonneault, M. H. Fast star, slow Star; old Star, young Star: subgiant rotation as a population and stellar physics diagnostic. *Astrophys. J.* **776**, 67–87 (2013).
- Hartman, J. D. *et al.* Deep MMT transit survey of the open cluster M37. III. Stellar rotation at 550 Myr. *Astrophys. J.* **691**, 342–364 (2009).
- Metcalfe, T. S. *et al.* Asteroseismology of the solar analogs 16 Cyg A and B from *Kepler* observations. *Astrophys. J.* **748**, L10–L16 (2012).
- Aigrain, S. *et al.* Testing the recovery of stellar rotation signals from *Kepler* light curves using a blind hare-and-hounds exercise. *Mon. Not. R. Astron. Soc.* **450**, 3211–3226 (2015).

18. Réville, V., Brun, A. S., Matt, S. P., Strugarek, A. & Pinto, R. F. The effect of magnetic topology on thermally driven wind: toward a general formulation of the braking law. *Astrophys. J.* **798**, 116–131 (2015).
19. Garraffo, C., Drake, J. J. & Cohen, O. The dependence of stellar mass and angular momentum losses on latitude and on active region and dipolar magnetic fields. *Astrophys. J.* **807**, L6 (2015).
20. Cohen, O., Drake, J. J., Kashyap, V. L. & Gombosi, T. I. The effect of magnetic spots on stellar winds and angular momentum loss. *Astrophys. J.* **699**, 1501–1510 (2009).
21. Brown, T. The metastable dynamo model of stellar rotational evolution. *Astrophys. J.* **789**, 101–113 (2014).
22. Morin, J., Dormy, E., Schriener, M. & Donati, J.-F. Weak- and strong-field dynamos: from the Earth to the stars. *Mon. Not. R. Astron. Soc.* **418**, L133–L137 (2011).

Acknowledgements We wish to thank the Kavli Institute for Theoretical Physics and the organizers of the ‘Galactic Archaeology and Precision Stellar Astrophysics’ program held from January to April 2015. We thank B. Shappee for useful discussions. J.L.v.S. is a Carnegie-Princeton Fellow. T.S.M. acknowledges the adopt-a-star crowdfunding program administered by White Dwarf Research Corp. G.R.D. acknowledges the support of the UK Science and Technology Facilities Council (STFC). M.H.P., T.S.M. and S.M. acknowledge support from NASA grant NNX15AF13G and National Science Foundation (NSF) grant AST-1411685. T.C. and R.A.G. received grants from the Centre National d’Études Spatiales (CNES) at the French Alternative Energies and Atomic Energy Commission (CEA). R.A.G. also acknowledges funding from the

European Community Seventh Framework Programme (FP7/2007–2013) under grants 312844 (SPACEINN) and 269194 (IRSES/ASK). This research was supported in part by the NSF under grant NSF PHY11-25915. Funding for the Stellar Astrophysics Centre is provided by The Danish National Research Foundation (grant agreement DNRF106). This research is supported by the ASTERISK project (ASTERoseismic Investigations with SONG and Kepler) funded by the European Research Council (grant agreement 267864). V.S.A. acknowledges support from VILLUM FONDEN (research grant 10118).

Author Contributions J.L.v.S. provided the physical interpretation of the rapid rotation and the rotational and stellar modelling. M.H.P. contributed to the design of the rotation tracer code, to ongoing YREC (Yale rotating stellar evolution code) development, and to interpretation of the results. T.S.M. carried out asteroseismic modelling portal (AMP) modelling of all targets. V.S.A. carried out Bayesian stellar algorithm (BASTA)–Garching stellar evolution code modelling. T.C., R.A.G. and S.M. developed and implemented the analysis of rotational modulation in Kepler data. G.R.D. provided updated asteroseismic mode frequencies used in the modelling.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.L.v.S. (jvansaders@obs.carnegiescience.edu) or R.A.G. (rafael.garcia@cea.fr) for computers code.

METHODS

Sample selection. Our sample can be divided into two principal target types: Kepler Asteroseismic Science Consortium (KASC) targets, and Kepler Objects of Interest (KOIs). We focus on those stars with (modelled) ZAMS T_{eff} values (defined as the point at which hydrogen fusion dominates the stellar luminosity) below 6,200 K, where magnetic braking should be most important. We show the positions of the selected stars on a Hertzsprung–Russell diagram in Extended Data Fig. 1, and period–age plot in Extended Data Fig. 2.

As described elsewhere⁹, the asteroseismic sample is drawn from a magnitude-limited sample of 2,000 Sun-like stars that were selected for a one-month period of short-cadence (~ 1 -minute) Kepler observations on the basis of their properties in the Kepler Input Catalog (KIC). Of these stars, roughly 500 displayed evidence of solar-like oscillations. A subset of targets with detections of oscillations that show high signal to noise ratio detections of oscillations were selected for continued monitoring over Kepler quarters 5–17. Of this sample, the mode frequencies for a subset of 61 high-signal-to-noise stars were extracted; there are high-resolution spectroscopic data for 46 of these. We modelled 42 of these 46 stars with the asteroseismic modelling portal (AMP, described below), excluding 4 targets whose spectra contained a complicated pattern of mixed modes. Of the 42 modelled targets, 11 were both detected in spot modulation and classified as ‘simple’ solar-like oscillators that did not show the seismic hallmarks of F-stars and evolved subgiants. A further three (non-overlapping) targets were added⁷. Of this sample of 14, 12 targets have AMP ZAMS T_{eff} values of less than 6,200 K, yielding a total of 12 stars in the KASC sample.

The KOI sample⁹ was selected from the 77 KOIs observed in short cadence that displayed signatures of solar-like oscillations. Of these, 35 power spectra were of sufficient quality to extract individual mode frequencies to be modelled, 33 of which represent unevolved main-sequence stars. A subset of 11 have periods detected via spot modulation⁶, 7 of which have an AMP ZAMS T_{eff} of less than 6,200 K.

Finally, we add the two well studied stars from the 16 Cygni binary to our sample; for these stars, asteroseismic ages¹⁶ and rotation periods have been inferred from asteroseismic mode splittings²³. In total, 21 stars are addressed in this analysis. Where available, we use the updated asteroseismic frequencies of ref. 24. Extended Data Table 1 shows the seismic (mass, age) and spectroscopic (T_{eff} , [Fe/H]) values and rotation periods for these stars.

Age and period measurements. Asteroseismic ages are determined using two methodologies: AMP, which provides the ages used in most of this paper; and the Bayesian stellar algorithm (BASTA) pipeline, used to verify that the discrepancies in predicted and observed rotation periods are not the result of pipeline choice. AMP uses a genetic algorithm to perform a search for the global χ^2 minimum between the stellar observables and stellar model values⁹. The algorithm uses the Aarhus stellar evolution code (ASTEC) and adiabatic pulsation code (ADIPLS) to compute oscillation frequencies. The BASTA pipeline uses a Bayesian approach to model stars with a grid of models produced with the Garching stellar evolution code (GARSTEC). The input physics of the stellar models used in each method are detailed in refs 8–10.

Both methods use frequency spacings and spectroscopic constraints to identify the optimal stellar properties, but AMP also uses the individual frequencies by employing an empirical correction for surface effects. There are two main differences between the models used by BASTA and those used by AMP. BASTA–GARSTEC uses a fixed relationship between the initial helium and metallicity, anchored to zero metallicity at the primordial helium abundance and assuming $\Delta Y/\Delta Z = 1.4$ to reproduce the solar values (Y is the mass fraction of helium and Z is the mass fraction of all other elements excluding hydrogen and helium). It also uses a single solar-calibrated value of the mixing-length parameter for all models. AMP–ASTEC allows the initial helium to float independently of metallicity, and searches a wide range of values for the mixing-length parameter. Both sets of models include diffusion, although BASTA–GARSTEC includes both helium and heavy-metal diffusion, while AMP–ASTEC considers only helium diffusion.

We extract rotation periods using techniques⁵ that we summarize briefly here (full period-extraction diagrams are available at http://irfu.cea.fr/Phocear/Vie_des_labos/Ast/ast_techique.php?id_ast=3607). For the corrected light curve of each Kepler star, the autocorrelation function (ACF) and a wavelets decomposition (period-time) are calculated. We collapse the wavelet decomposition on the period axis to obtain the global wavelet power spectrum (GWPS), and the peaks of this GWPS are fitted with gaussian functions. In parallel, we identify the peaks of the ACF. The derived surface rotation period is the result of the comparison of the ACF and GWPS analyses and is confirmed by a visual inspection of the light curves.

Stellar rotation models. We use a theoretical model grid¹⁴ (using OPAL rather than OP opacities; all other inputs are unchanged), utilize the same loss-law calibration and form as in ref. 14, and assume solid-body rotation. The model grid is expanded to cover a wider range of metallicities and masses, namely $[Z/H] = -0.4$ to $[Z/H] = +0.4$, assuming a helium enrichment of $\Delta Y/\Delta Z = 1.0$ and no diffusion

or gravitational settling. We use the ‘fast launch’ conditions¹⁴ for modelling the rotation, but have validated that our results are insensitive to the choice of initial conditions. Changing the launch conditions typically shifts the period ratio (in sense of expected/observed) by less than 50% of the quoted errors, and shifts the fitted critical Rossby number to $\text{Ro}_{\text{crit}} = 2.15 \pm 0.08$. The model τ_{cz} is the local convective overturn timescale, defined as the ratio of the typical mixing length to the convective velocity at one pressure scale height above the base of the convective envelope in the mixing length theory of convection. Under this definition, the solar rotation period (P_0) is 25.4 days, $\tau_{\text{cz},\odot}$ is 1.015×10^6 s, and $\text{Ro}_{\odot} = 2.16$.

The weakened magnetic braking is modelled by modifying the braking law such that a star with $P/\tau_{\text{cz}} > \text{Ro}_{\text{crit}}$ is evolved under the assumption of conservation of angular momentum, such that the rotation period depends only on the changing moment of inertia of the star as it evolves. The modified loss law is given by the following equations (based on equations (1) and (2) in ref. 14):

$$\frac{dJ}{dt} = \begin{cases} f_K K_M \omega \left(\frac{\omega_{\text{crit}}}{\omega_{\odot}} \right)^2, & \omega_{\text{crit}} \leq \frac{\omega_{\text{cz}}}{\tau_{\odot}}, \text{ Ro} \leq \text{Ro}_{\text{crit}} \\ f_K K_M \omega \left(\frac{\omega_{\text{cz}}}{\omega_{\odot} \tau_{\text{cz},\odot}} \right)^2, & \omega_{\text{crit}} > \frac{\omega_{\text{cz}}}{\tau_{\odot}}, \text{ Ro} \leq \text{Ro}_{\text{crit}} \\ 0, & \text{Ro} > \text{Ro}_{\text{crit}} \end{cases}$$

$$\frac{K_M}{K_{M,\odot}} = c(\omega) \left(\frac{R}{R_{\odot}} \right)^{3.1} \left(\frac{M}{M_{\odot}} \right)^{-0.22} \left(\frac{L}{L_{\odot}} \right)^{0.56} \left(\frac{P_{\text{phot}}}{P_{\text{phot},\odot}} \right)^{0.44}$$

where f_K is a constant factor used to scale the loss law during the empirical fitting; ω_{crit} is the saturation threshold (important only at young ages); τ_{cz} is the convective overturn timescale; P_{phot} is the pressure at the photosphere; R is the radius; M is the mass; L is the luminosity; and \odot refers to the Sun. The term $c(\omega)$ sets the centrifugal correction; because our stars are slowly rotating and the correction should be small, we set $c(\omega)$ to a constant value of 1. This braking law is fit to open-cluster data and the Sun, where the initial rotation period, disk-locking timescale, ω_{crit} , and f_K were allowed to vary, and all other parameters were determined using stellar evolutionary models¹⁴. When fitting for an optimal Ro_{crit} , we keep the parameters of the magnetic braking law calibrated on the Sun and open clusters fixed, and vary only the Ro_{crit} at which braking is allowed to cease. Ro_{crit} is optimized using a χ^2 figure of merit (valid under the assumption of independent observations and Gaussian uncertainties): $\chi^2 = \sum_i (P_{\text{obs},i} - \text{mod}_i)^2 / (\sigma_{\text{obs},i}^2 + \sigma_{\text{mod},i}^2)$, where $\sigma_{\text{obs},i}$ is the observational uncertainty on the extracted period, and $\sigma_{\text{mod},i}$ represents the uncertainty on the model period given the uncertainties on the input masses, ages, and compositions. We derive uncertainties on Ro_{crit} using bootstrap resampling, drawing a 21-star sample with replacement from the original data 50,000 times, and recalculating the best-fit Ro_{crit} for each realization. Cluster data and the Sun are not used in this fit. An alternate fit allowing parameters important for late-time braking to vary (f_K , Ro_{crit}) and including intermediate-age and older rotation data from the seismic sample, NGC 6819, and the Sun (52 stars in total) yields a best-fit Ro_{crit} of 2.1 ± 0.1 , with $f_K = 8.4 \pm 0.2$.

Predicted model periods are obtained by using the mass and age from the asteroseismic pipelines coupled with the spectroscopic metallicity^{8–10,16}. Model uncertainties are estimated by generating 50,000 (20,000 for Ro_{crit} + f_K fit) realizations of the input parameters (M , t and [Fe/H]), where values are drawn from a Gaussian distribution centred on the observed value, with 1σ errors defined by the observational uncertainties. While we search in the fundamental space of mass, age and composition, we select only models which fall within 5σ of the observed T_{eff} . This constraint has little or no effect for unevolved stars, but ensures that stars at the turnoff (KIC 6196457 and KIC 8349582 in particular) are not assigned artificially long rotation periods due to mass–age combinations that fall on the subgiant branch. 1σ uncertainties on the model periods are defined as the values that enclose 68% of the resulting models.

Empirical gyrochronology relations. We verify that the unexpectedly rapid rotation in old, solar-like stars is independent of the spin-down prescription by repeating our exercise with an empirical literature gyrochronology relation². We replicate Fig. 2 in Extended Data Fig. 3 with predicted periods drawn from an empirical gyrochronology calibration, based on equation (32) in ref. 2:

$$t = \frac{\tau}{k_C} \ln \left(\frac{P}{P_0} \right) + \frac{k_1}{2\tau} (P^2 - P_0^2)$$

where t is the age, τ is the convective overturn timescale, P is the period, and P_0 is the initial period. We adopt values for k_C of 0.646 million years (Myr) per day and for k_1 of 452 days per Myr; $P_0 = 1.1$ days, and the global $\tau - T_{\text{eff}}$ relation is as used in refs 2,4. 50,000 realizations of the combination (T_{eff} , t) are drawn from a Gaussian

distribution centred on the measured values, with a 1σ width defined by the quoted observational errors on the central values. These empirical relationships do not account for the physical expansion of stars as they evolve (particularly near the end of the main sequence) and therefore tend to predict somewhat more rapid rotation than do full theoretical models near the main sequence turnoff.

Cluster data. To provide comparison with the typical gyrochronological calibrators, we draw cluster data from a variety of literature sources. For the cluster M37 we adopt the following cluster parameters¹⁵: extinction, $E(B-V) = 0.227 \pm 0.038$ mag; total metal abundance $[M/H] = 0.045 \pm 0.044$ dex; age = 550 ± 30 Myr. Rotation data and cluster parameters²⁵ for Praesepe (M44) are included, with $E(B-V) = 0.027 \pm 0.004$, $[Fe/H] = 0.11 \pm 0.03$, and $\log(\text{age}) = 8.77 \pm 0.1$. For NGC 6811, we adopt the $g-r$ colours, $E(B-V)$ value of 0.1, and rotation periods as in ref. 26, as well as the $[M/H]$ value of -0.1 ± 0.01 and age of 1.00 ± 0.05 Gyr from ref. 27. Finally, for NGC 6819 we use the rotation periods and $B-V$ colours from ref. 4, with the age (2.5 ± 0.2 Gyr) and adopted metallicity (0.09 ± 0.03) from ref. 28. $B-V$ colours are converted into temperatures and stellar masses using Yale rotating stellar evolution (YREC) isochrones²⁹. We model cluster stars in the same manner as the seismic targets, with 10,000 mass-age-composition realizations for each star. We display the mean cluster rotation periods for all stars within the ZAMS T_{eff} bins, with errors representing the 16th and 84th percentiles. In M37 and Praesepe in particular, the rotational distribution displays a range resulting from spread in the initial rotation periods.

Sample biases. We demonstrate that our results are unlikely to be a consequence of selection bias in our sample. The sample is subject to two sources of selection bias: asteroseismic detectability, and the detectability of spot modulation.

Detailed asteroseismic analysis requires a high signal-to-noise detection of the power excess from oscillations. Oscillation amplitudes scale roughly as $A_{\text{max}} \propto (L/M) (T_{\text{eff}})^{-2}$ (equation (7) from ref. 30, referring to $l=0$ radial modes); seismic samples are therefore strongly biased towards more-massive stars. There is also a bias towards bright targets, where lower noise levels contribute to detectability. Our sample is drawn from two subsets of stars: the one-month survey stars from the seismic sample, and the KOIs. We expect the one-month survey seismic detections at magnitudes of K_p of less than ~ 10 , while the roughly 1,000-day time series in short cadence collected for the KOI sample allow detections out to $K_p \approx 12$, which well describes the actual magnitude distribution of our sample (see Fig. 6 in ref. 30). The strong trends with magnitude and mass are well predicted by basic scaling arguments, save for the dependence on activity: active stars are less likely to be detected in oscillations³⁰. Our sample is selected seismically, and we do not expect the well understood seismic biases to favour rapid rotators (apart from the obvious mass dependence).

Variability owing to starspots scales with the rotation period, in the sense that more rapid rotation is associated with higher amplitudes of variability¹⁵. One could imagine that we are detecting the rapidly rotating tail of a distribution of rotation periods, or detecting objects spun up by binary/planetary interactions or mergers.

This first case is at odds with what we know from open clusters: as late as 2.5 Gyr, there is a converged, well defined rotational sequence that shows very little scatter at fixed mass⁴. If we are in fact detecting a rapidly rotating subset of the population, the dispersion in rotation and spin-down rates must set in after several billion years, or it would be visible in the open-cluster data. If there is dispersion in the rotation periods, it represents a serious challenge to the validity of gyrochronology for old stars, regardless of its source.

The pipeline used to extract the rotation periods for this work has been tested with an injection and recovery exercise¹⁷. Our recovery fraction is shown in Extended Data Fig. 4, and demonstrates that we should be able to detect stars that are substantially less active than the Sun at longer periods. However, this exercise does not account for stars that simply cease to have spots to detect on their surfaces; under this scenario, slow rotators could exist but be undetectable. We cannot directly combat this concern given our current data set, although we can examine the case of 16 Cygni (16 Cyg). 16 Cyg A and B are not detected in spot modulation; their periods are derived from asteroseismic frequency splittings, which yield periods that probe the envelope rotation²³. If we assume that these stars have solar-like rotation profiles, then the seismic rotation periods are directly comparable to the surface periods. This pair displays the same anomalously rapid rotation as objects detected in spot modulation, providing evidence against the argument that stars undetected in spot modulation are simply more slowly rotating. It is also worth noting that our own Sun would be undetectable during the minimum of its activity cycle (see ref. 17). Our non-detections could equally be the result of the normal variations in the activity of Sun-like stars, rather than a period bias.

Finally, we examine the possibility of interactions or mergers with other bodies. In our sample, 16 Cyg A and B, and KIC 3427720 and KIC 9139151, are

known or suspected binaries⁵. In each case, the components are well separated, and the binary orbits are estimated well in excess of 10,000 years. In order for a companion to affect rotation considerably, it must be at orbital periods comparable to the rotation period, and will therefore be unresolved. The KOI sample has undergone the extensive vetting that is associated with planet detection; all planets are confirmed, and there is no evidence of transit timing variations that would accompany a close stellar companion. System stability is unlikely for binary orbits of the order of 30 days that contain even a low-mass stellar companion³¹. Likewise, there is no evidence for interaction between the planets and the host stars in the KOI sample⁶, and no known hot Jupiters. In the case of the seismic sample, there is no evidence for double-lined binaries, photometric-spectroscopic temperature disagreements, multiple oscillating components, or unusual dilution of the seismic power spectra, and no evidence of eclipses. Finally, if mergers (planetary or stellar) were responsible for all detections of rapid rotation, then the 50% detection rate of the 'simple stars'⁷ in spot modulation implies an uncomfortably high merger rate.

The asteroseismic age scale. We carry out two tests to demonstrate that the discrepancy between the expected and observed rotation periods is not due to a systematic bias in the ages with roots in the asteroseismic age scale. We show that ages derived with the BASTA pipeline display the same trend in rotation period, and that systematically shifting the asteroseismic ages, while improving the fit, is inferior to instituting a Rossby threshold.

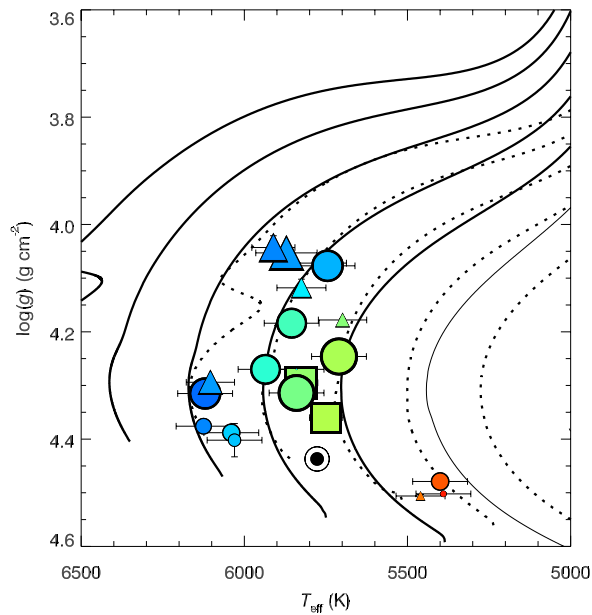
Extended Data Fig. 5 provides period ratio plots using the BASTA ages and BASTA ZAMS T_{eff} determinations. The systematic trend in the period ratios survives. The Barnes relation² fits with $\chi^2 = 184.3$, and the fiducial models¹⁴ with $\chi^2 = 68.4$. A fit for Ro_{crit} using the BASTA ages yields $\text{Ro}_{\text{crit}} = 2.67 \pm 0.50$. Bootstrap resampling demonstrates that this number is sensitive to whether KIC 8349582 is drawn; if KIC 8349582 is excluded, the fit becomes $\text{Ro}_{\text{crit}} = 2.12 \pm 0.12$.

We also investigate the possibility that the seismic age scale is systematically shifted relative to the true ages. We perform model fits with the fiducial braking law with an extra parameter that allows for a systematic age shift. For the AMP ages, χ^2 is minimized with the Barnes relation with a systematic shift of 35% and a χ^2 of 78.5. Likewise, the fiducial models¹⁴ prefer a shift of $20 \pm 3\%$ with a χ^2 of 26.9. In both cases, the required systematic shifts are larger than the estimated 9.6% systematic uncertainties in seismic ages¹⁰.

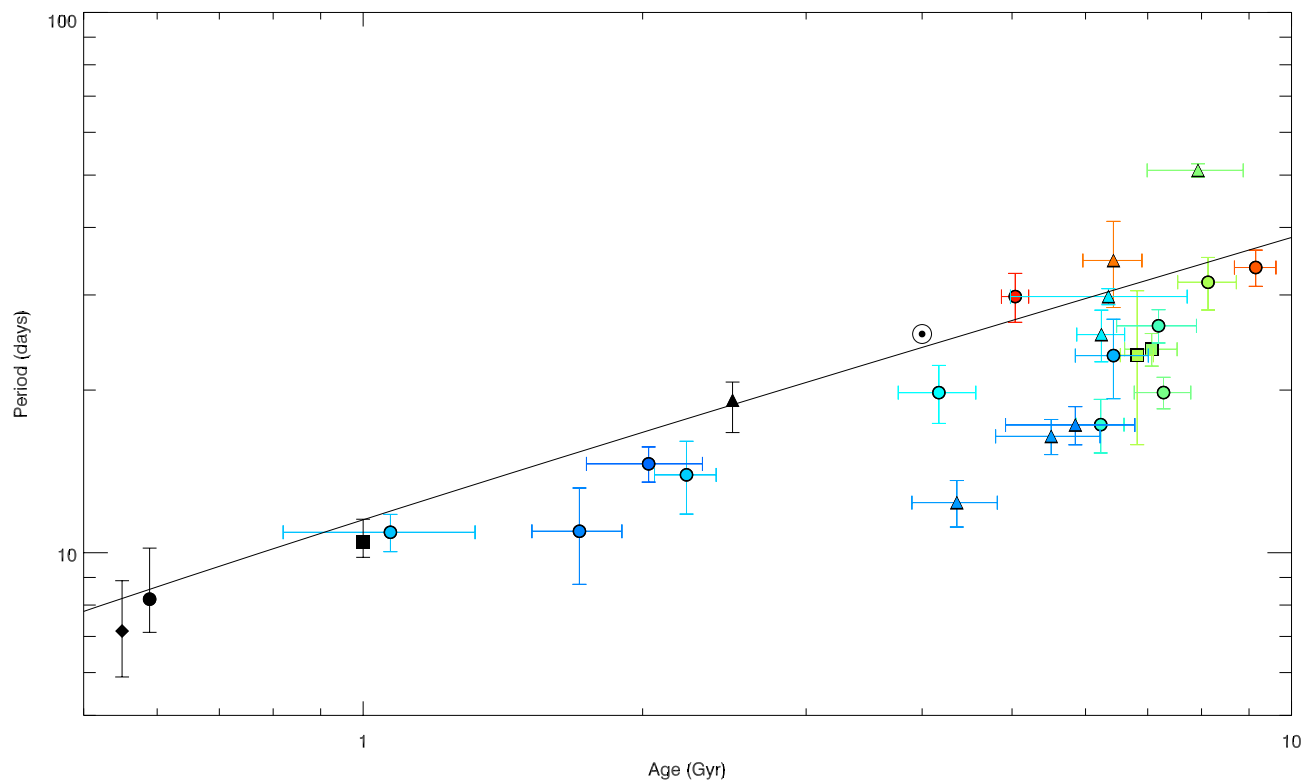
Finally, to verify that we are not biased by the fact that the ages and periods were determined using different evolution codes, we tune the physics in the fiducial models¹⁴ to match that of the AMP models, and predict the rotation periods for the central AMP values of the masses, ages, and compositions of each star. In particular, we match the diffusion physics, opacity tables, equation of state, helium and metal abundances, boundary conditions, and important nuclear reaction rates present in the ASTEC code used for AMP. The results are presented in Extended Data Fig. 6, and demonstrate that the discrepancy between the predicted and observed periods is preserved. We conclude that our result is not the consequence of assumptions about the stellar physics included in models.

Code availability. The AMP science code used to infer stellar ages can be downloaded at <https://amp.phys.au.dk/about/evolpack>. Code for the period extraction and rotational evolution will be publicly released upon completion of the necessary documentation. YREC likewise has no public documentation, and has not been publicly released. BASTA is undergoing major revisions for increased speed and is not yet publicly available.

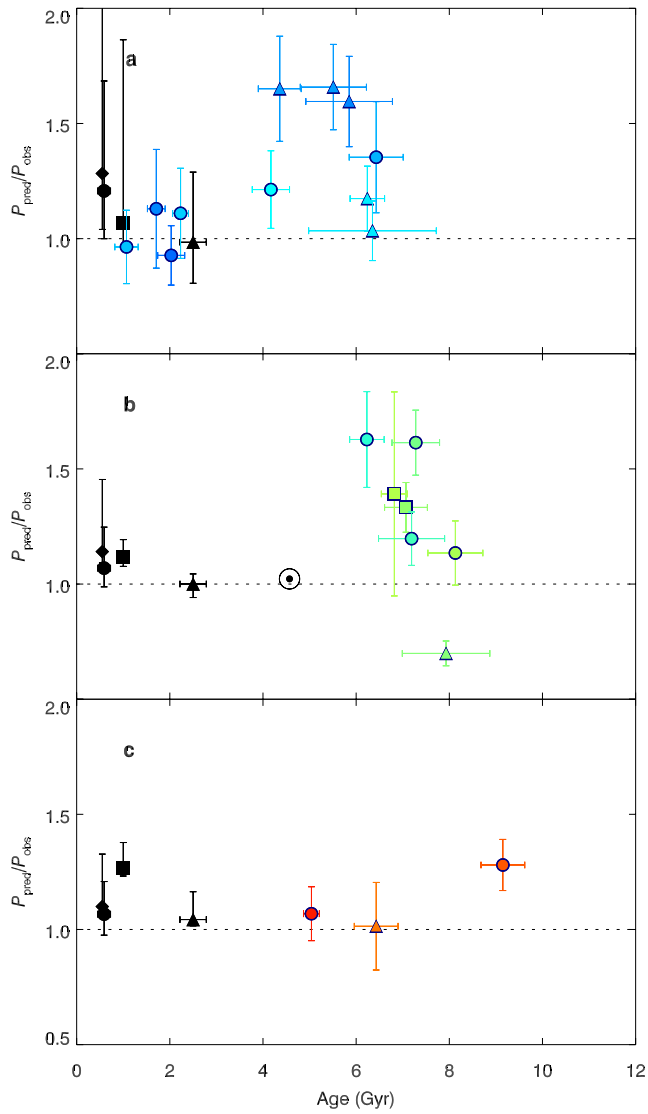
23. Davies, G. R. *et al.* Asteroseismic inference on rotation, gyrochronology and planetary dynamics of 16 Cygni. *Mon. Not. R. Astron. Soc.* **446**, 2959–2966 (2015).
24. Davies, G. R. *et al.* Oscillation frequencies for 35 Kepler solar-type planet-hosting stars using Bayesian techniques and machine learning. Preprint at <http://arXiv.org/abs/1511.02105> (2015).
25. Kovacs, G. *et al.* Stellar rotational periods in the planet hosting open cluster Praesepe. *Mon. Not. R. Astron. Soc.* **442**, 2081–2093 (2014).
26. Meibom, S. *et al.* The Kepler cluster study: stellar rotation in NGC6811. *Astrophys. J.* **733**, L9–L14 (2011).
27. Yontan, T. CCD UVRI photometry of NGC 6811. *Astrophys. Space Sci.* **355**, 267–281 (2015).
28. Jeffries, M. W. WCSS 40007: a detached eclipsing binary near the turnoff of the open cluster NGC 6819. *Astron. J.* **146**, 58–78 (2013).
29. An, D. *et al.* The distances to open clusters from main-sequence fitting. III. Improved accuracy with empirically calibrated isochrones. *Astron. J.* **655**, 233–260 (2007).
30. Chaplin, W. J. *et al.* Predicting the detectability of oscillations in solar-type stars observed by Kepler. *Astrophys. J.* **732**, 54–62 (2011).
31. Holman, M. J. & Weigert, P. A. Long-term stability of planets in binary systems. *Astron. J.* **117**, 621–628 (1999).



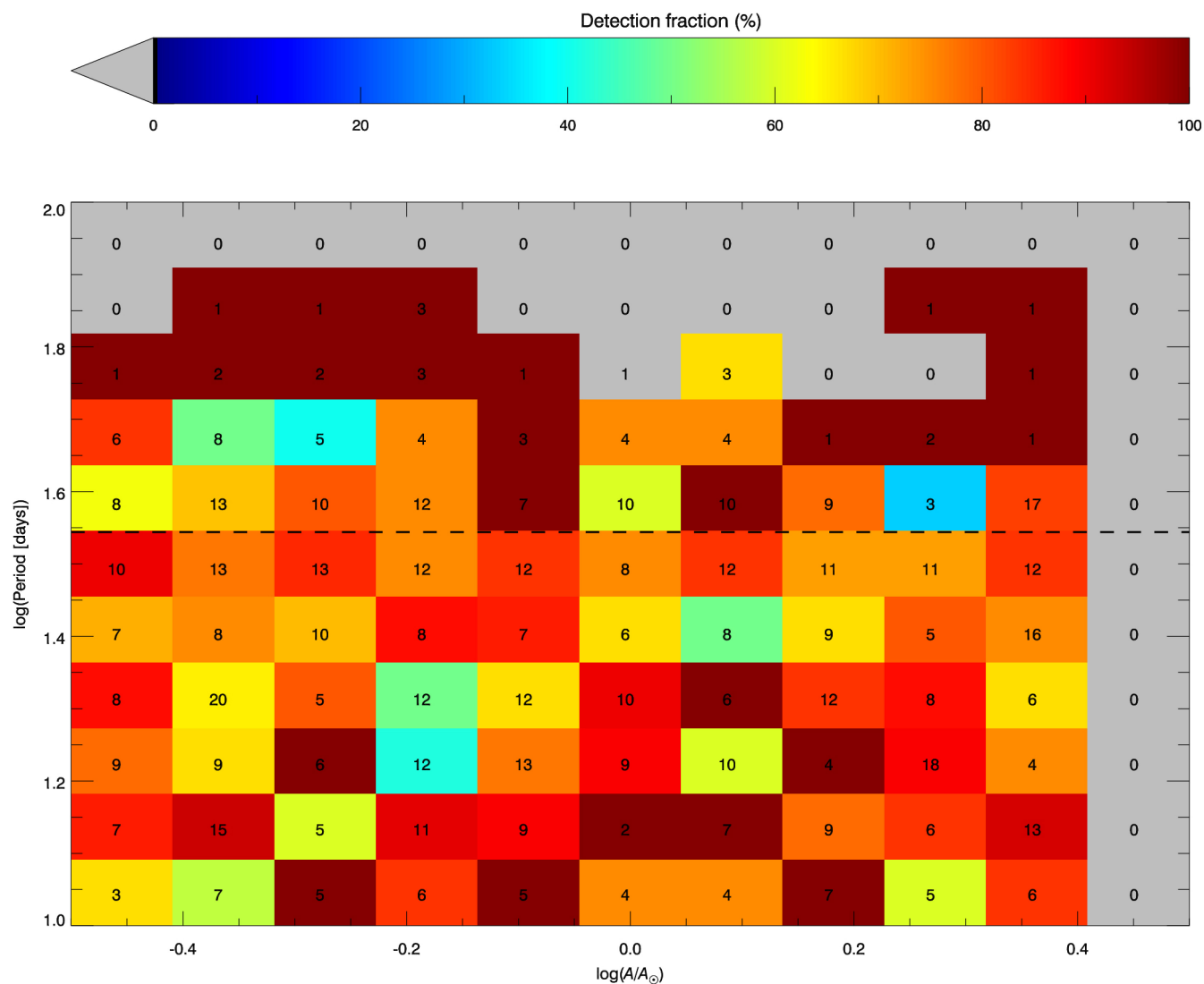
Extended Data Figure 1 | The positions of all 21 Kepler stars on the Hertzsprung–Russell diagram. We plot spectroscopic T_{eff} (a proxy for mass) versus seismic $\log(g)$ (surface gravity), with 1σ observational error bars; the symbol size is proportional to the period ratio (AMP ages, fiducial models¹⁴). Colours and symbol conventions are as in Fig. 2. Evolutionary tracks are overplotted for $[Z/H] = +0.3$ (dotted lines) and $[Z/H] = -0.1$ (solid lines), for masses $0.8\text{--}1.3 M_{\odot}$ in increments of $0.1 M_{\odot}$. ($[Z/H] = +0.3$, $M = 0.8 M_{\odot}$ is beyond the plotted area.)



Extended Data Figure 2 | Period–age plot of sample stars. The 21-star sample, with observed rotational periods plotted against AMP asteroseismic ages. Symbol conventions are as in Fig. 2. The solid line denotes the empirical relation² for $T_{\text{eff}} = 5,800 \text{ K}$ (approximately equal to the mean sample T_{eff}). All error bars represent 1σ .



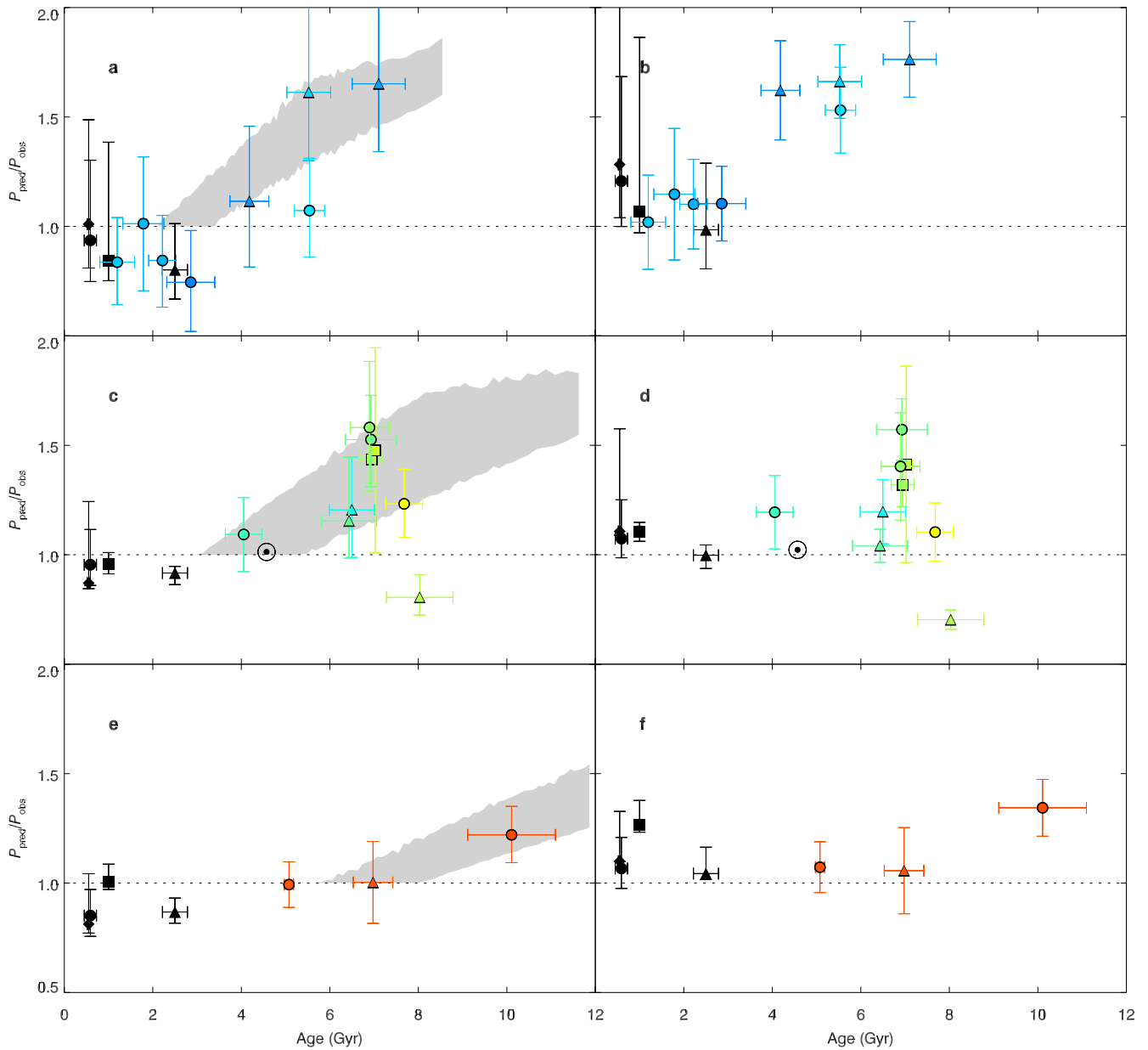
Extended Data Figure 3 | Period ratios using empirical gyrochronology relations. Ratios of predicted periods² to observed periods are plotted as a function of the AMP asteroseismic age, and divided according to AMP ZAMS T_{eff} (a, ZAMS $T_{\text{eff}} = 5,900\text{--}6,200\text{ K}$; b, ZAMS $T_{\text{eff}} = 5,600\text{--}5,900\text{ K}$; c, ZAMS $T_{\text{eff}} = 5,100\text{--}5,400\text{ K}$.) Error bars represent 1σ . Symbol conventions are as in Fig. 2.



Extended Data Figure 4 | Detectability of stars in spot modulation.

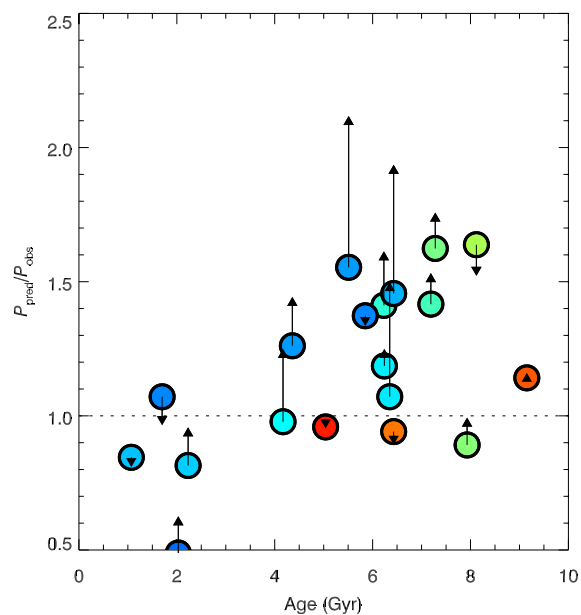
Detection fractions for the 750 stars with noise in the hound-and-hare exercise of ref. 17, as a function of activity level A (where the activity level of the Sun is defined as $A_{\odot} = 1$) and rotation period. The total number of

light curves searched for periodicity in each cell is overplotted. The dashed black line at $P = 35$ days represents the expected period for stars like the Sun under traditional gyrochronology relations found in the literature.



Extended Data Figure 5 | Predicted versus observed rotation periods using ages determined with BASTA. a, c, e, Plotted are the ratios of the periods predicted using the fiducial models¹⁴ to the observed rotation periods, as a function of stellar age. The grey band represents the offset expected from models in which $Ro_{\text{crit}} = 2.16$. All error bars represent 1σ .

b, d, f, Ratios of the predicted periods obtained from the empirical relation² to the observed periods, plotted against stellar age. Stars are divided according to ZAMS T_{eff} , using BASTA ZAMS T_{eff} values: a, b, 5,900–6,200 K; c, d, 5,500–5,900 K; e, f, 5,100–5,400 K. All symbol conventions are as in Fig. 2.



Extended Data Figure 6 | The shift in the period ratios induced by changing the stellar model input physics. Circles are colour-coded according to ZAMS T_{eff} as in Fig. 2. Ratios of the periodicity expected from the fiducial model¹⁴ to the observed periodicity are plotted against age. Arrows denote the shift in the period ratio that occurs when YREC models¹⁴ are run to match the AMP-ASTEC physics.

Extended Data Table 1 | Rotation periods, asteroseismic data and spectroscopic quantities for sample stars

KIC	AMP				BASTA/GARSTEC			Spectroscopic		Period	Note
	Mass	Age	log(g)	ZAMS T_{eff}	Mass	Age	ZAMS T_{eff}	T_{eff}	[Fe/H]		
16Cyg A	1.10±0.02	7.07±0.46	4.295	5677	1.04±0.01	6.95±0.26	5668	5825±50	+0.09±0.02	23.8±1.7	seismic period
16Cyg B	1.06±0.02	6.82±0.28	4.360	5629	0.998±0.005	7.02±0.14	5592	5750±50	+0.05±0.02	23.2±7.4	seismic period
3427720	1.13±0.04	2.23±0.17	4.388	5985	1.12±0.02	2.22±0.31	6019	6040±84	-0.03±0.09	13.9±2.1	seismic
3656476	1.17±0.03	8.13±0.59	4.246	5642	1.07±0.01	7.68±0.42	5525	5710±84	+0.25±0.09	31.7±3.5	seismic
5184732	1.27±0.04	4.17±0.40	4.270	5905	1.18±0.02	4.05±0.42	5810	5840±84	+0.38±0.09	19.8±2.4	seismic
6116048	1.01±0.03	6.23±0.37	4.270	5838	1.06±0.02	5.54±0.34	5943	5935±84	-0.24±0.09	17.3±2.0	seismic
6196457	1.23±0.04	5.51±0.71	4.053	6064	1.21±0.02	5.52±0.50	5991	5871±94	+0.17±0.11	16.4±1.2	KOI
6521045	1.04±0.02	6.24±0.37	4.118	5933	1.11±0.02	6.50±0.51	5886	5825±75	+0.02±0.10	25.3±2.8	KOI
7680114	1.13±0.03	7.19±0.70	4.184	5801	-- --	-- --	-- --	5855±84	+0.11±0.09	26.3±1.9	seismic
7871531	0.84±0.02	9.15±0.47	4.479	5253	0.84±0.02	10.10±0.99	5240	5400±84	-0.24±0.09	33.7±2.6	seismic
8006161	1.04±0.02	5.04±0.17	4.502	5165	0.948±0.005	5.08±0.10	5250	5390±84	+0.34±0.09	29.8±3.1	seismic
8349582	1.19±0.04	7.93±0.94	4.178	5695	1.07±0.02	8.03±0.75	5630	5699±74	+0.30±0.10	51.0±1.5	KOI
9098294	1.00±0.03	7.28±0.51	4.314	5718	1.01±0.02	6.93±0.57	5734	5840±84	-0.13±0.09	19.8±1.3	seismic
9139151	1.14±0.03	1.71±0.19	4.376	6092	1.16±0.02	1.79±0.46	6019	6125±84	+0.11±0.09	11.0±2.2	seismic
9955598	0.96±0.01	6.43±0.47	4.506	5307	0.89±0.01	6.98±0.45	5250	5460±75	+0.08±0.10	34.7±6.3	KOI
10454113	1.19±0.04	2.03±0.29	4.315	6138	1.15±0.03	2.86±0.54	6095	6120±84	-0.06±0.09	14.6±1.1	seismic
10586004	1.16±0.05	6.35±1.37	4.072	5943	1.18±0.03	6.43±0.62	5753	5770±83	+0.29±0.10	29.8±1.0	KOI
10644253	1.13±0.05	1.07±0.25	4.402	6001	1.16±0.02	1.20±0.39	5991	6030±84	+0.12±0.09	10.91±0.87	seismic
10963065	1.07±0.02	4.36±0.46	4.294	6063	1.09±0.02	4.18±0.44	6076	6104±74	-0.20±0.10	12.4±1.2	KOI
11244118	1.10±0.05	6.43±0.58	4.077	6023	1.13±0.02	6.90±0.44	5677	5745±84	+0.35±0.09	23.2±3.9	seismic
11401755	1.03±0.05	5.85±0.93	4.043	6094	1.06±0.03	7.10±0.60	6057	5911±66	-0.20±0.06	17.2±1.4	KOI

Units are as follows: mass, solar masses; age, Gyr; log(g), g cm^{-2} , T_{eff} , K; period, days. Quoted errors are given to 1σ .

Controlling many-body states by the electric-field effect in a two-dimensional material

L. J. Li^{1,2,3*}, E. C. T. O'Farrell^{1,2*}, K. P. Loh^{1,3}, G. Eda^{1,2,3}, B. Özyilmaz^{1,2} & A. H. Castro Neto^{1,2}

To understand the complex physics of a system with strong electron–electron interactions, the ideal is to control and monitor its properties while tuning an external electric field applied to the system (the electric-field effect). Indeed, complete electric-field control of many-body states in strongly correlated electron systems is fundamental to the next generation of condensed matter research and devices^{1–3}. However, the material must be thin enough to avoid shielding of the electric field in the bulk material. Two-dimensional materials do not experience electrical screening, and their charge-carrier density can be controlled by gating. Octahedral titanium diselenide (1T-TiSe₂) is a prototypical two-dimensional material that reveals a charge-density wave (CDW) and superconductivity in its phase diagram⁴, presenting several similarities with other layered systems such as copper oxides⁵, iron pnictides⁶, and crystals of rare-earth elements and actinide atoms⁷. By studying 1T-TiSe₂ single crystals with thicknesses of 10 nanometres or less, encapsulated in two-dimensional layers of hexagonal boron nitride, we achieve unprecedented control over the CDW transition temperature (tuned from 170 kelvin to 40 kelvin), and over the superconductivity transition temperature (tuned from a quantum critical point at 0 kelvin up to 3 kelvin). Electrically driving TiSe₂ over different ordered electronic phases allows us to study the details of the phase transitions between many-body states. Observations of periodic oscillations of magnetoresistance induced by the Little–Parks effect show that the appearance of superconductivity is directly correlated with the spatial texturing of the amplitude and phase of the superconductivity order parameter, corresponding to a two-dimensional matrix of superconductivity. We infer that this superconductivity matrix is supported by a matrix of incommensurate CDW states embedded in the commensurate CDW states. Our results show that spatially modulated electronic states are fundamental to the appearance of two-dimensional superconductivity.

The charge-carrier density—or equivalently, the Fermi energy—strongly controls phase transitions in correlated systems. Traditionally, charge-carrier density can be controlled by doping, that is, by chemical modification of the material. Unfortunately, the alteration of the system's chemical composition leads to the unavoidable introduction of disorder. In strongly correlated systems, owing to their exponential sensitivity to the local electronic environment, disorder can have a profound impact that masks the intrinsic many-body behaviour⁸. Hence, there is a growing need to change the charge-carrier density of strongly correlated systems without chemical means. The application of an electric field is one of the 'cleanest' ways (that is, it tends not to introduce disorder) to address many-body states because it is intrinsically homogeneous. However, electric fields are screened by the bulk material in three-dimensional metals, making their use difficult.

The Fermi energy not only controls the number of electric carriers (electrons or holes) but also the screening of external electric fields

and internal electron–electron interactions⁹. In two-dimensional (2D) systems the electrons move in a plane while the electric field propagates in three-dimensional space. Hence, 2D electrons are unable to screen electric fields, external or their own. Therefore, we chose to work with a 2D material, TiSe₂, of nanometre-scale thickness, and we used an ionic gel electrolyte gate to apply the electric field. In addition, the flake of TiSe₂ was encapsulated by a 2D dielectric, hexagonal boron nitride, to avoid external disorder and chemical oxidation and degradation caused by both air and the electrolyte.

Electrical transport measurements under electric-field-induced doping enabled us to construct the phase diagram shown in Fig. 1. Electron doping suppresses the CDW transition from 170 K to 40 K and superconductivity appears with a dome that peaks at 3 K. We show that the emergence of superconductivity is directly associated with the inhomogeneous electronic states that correspond to a periodic structure of the amplitude and phase shifts of the superconductivity order parameter. This periodic structure must be stabilized and pinned to the lattice, so we can infer the presence of an incommensurate CDW (ICDW) matrix surrounding commensurate CDW (CCDW) regions.

TiSe₂ nanosheets with thicknesses of 10 nm or less were prepared by mechanical exfoliation of a high-quality single crystal (Extended Data Fig. 1). The device fabrication and measurement details are described in Methods and Supplementary Information. In Fig. 2a we sketch the electric-field double layer transistor device used in our experiments; Fig. 2b shows a typical top-gate sweep at 285 K and the variation of the electron density as measured by the Hall effect (see Methods and Extended Data Fig. 2a, b). Using an electrolyte top gate and an electrostatic doped-Si bottom gate we could control the electron density up to about 10^{15} cm^{-2} and thereby explore the phase diagram of this 2D material.

Variation of the charge-carrier density n leads to strong variations of the sheet resistance R_S of the device, as shown in Fig. 2c, d. At low charge-carrier densities, one can clearly see a peak in the resistivity versus temperature. The CDW transition temperature¹⁰, T_{CDW} , corresponds to the inflection point of the resistance and was also measured by using the Hall effect (Extended Data Fig. 2c), to detect the reconstruction of the Fermi surface. On increasing the charge-carrier density, T_{CDW} decreases from 170 K to 40 K before becoming undetectable at around $n = 7.5 \times 10^{14} \text{ cm}^{-2}$.

On increasing the electron density we observe the superconductivity state, as shown in Fig. 2d. The superconductivity transition temperature, T_C , increases from 0 K at the quantum critical point (QCP) at $n = 1.2 \times 10^{14} \text{ cm}^{-2}$ up to approximately 3 K at an optimal density of $n = 7.5 \times 10^{14} \text{ cm}^{-2}$. We note that this is exactly the density at which the CDW signal vanishes, indicating a scenario of two competing orders¹¹. A further increase in density suppresses T_C , giving rise to the formation of a superconductivity dome, as shown in Fig. 1 together with representations of the inferred structure in each region of the phase diagram (discussed below).

¹Centre for Advanced 2D Materials and Graphene Research Centre, National University of Singapore, 117546, Singapore. ²Department of Physics, National University of Singapore, 117542, Singapore. ³Department of Chemistry, National University of Singapore, 117543, Singapore.

*These authors contributed equally to this work

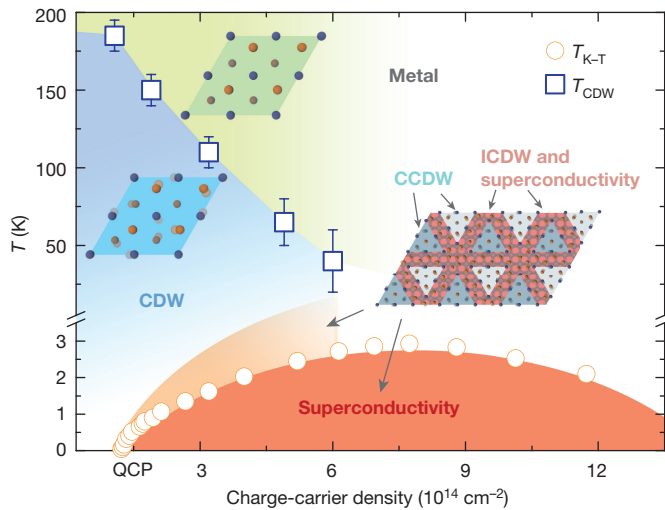


Figure 1 | Phase diagram of TiSe₂ under electron doping. Circles show T_{K-T} and squares show T_{CDW} . The insets show the lattice structure in each phase. In the CDW phases we illustrate the atomic displacements within an enlarged unit cell; in the phase in which ICDW and CDW coexist we schematically illustrate the ICDW domain walls between the CCDW regions as the red region (which we have exaggerated to occupy a single unit cell instead of a few nanometres). The error bars define the difference in T_{CDW} between the values derived from resistivity versus temperature curves and those derived from charge-carrier density versus temperature curves.

When the superconducting coherence length $\xi(T)$ becomes larger than the sample thickness close to T_C we expect the material to behave as a 2D system and the superconducting transition is anticipated to be of the Kosterlitz–Thouless (K–T) type with vortex–antivortex unbinding¹². One of the trademarks of the K–T transition is the broadening of the resistance with lowering of the temperature, as shown in Fig. 2d.

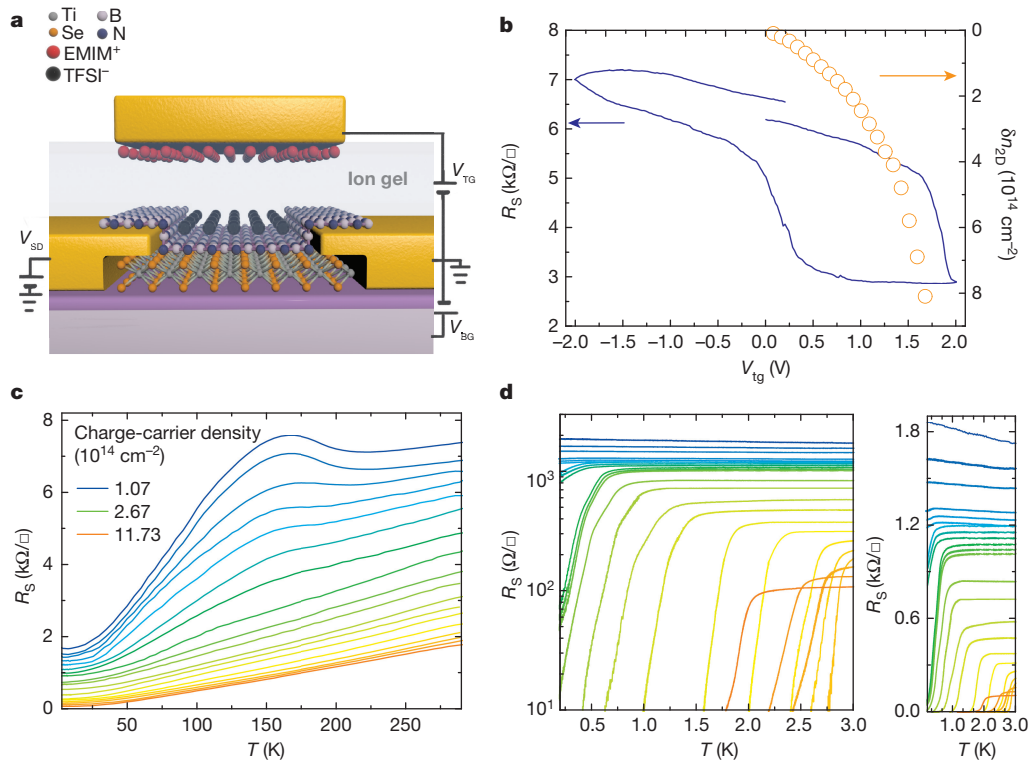


Figure 2 | Characterization of the field effect device and the resistance at different doping levels by gating. **a**, Sketch of the electric-field double layer transistor device. V_{SD} , source–drain voltage; V_{TG} , top gate voltage; V_{BG} , bottom gate voltage. **b**, Typical device characteristic under ion gel

For the K–T transition the resistivity is expected to scale with the coherence length as:

$$\xi(T) \approx a \exp(-b/\sqrt{T - T_{K-T}}) \quad (1)$$

where T_{K-T} is the K–T transition temperature, and a and b are material parameters. The experimental result reproduces this relation close to T_{K-T} , as shown in Fig. 3a for a charge-carrier density of $n = 2.67 \times 10^{14} \text{ cm}^{-2}$. We also observe current–voltage scaling¹³ in the superconductivity phase ($V \propto I^\alpha$) with $\alpha = 5$ for $n = 5.9 \times 10^{14} \text{ cm}^{-2}$ at the lowest temperature (Extended Data Fig. 3a). By fitting to equation (1), the K–T transition temperatures can be extracted for each doping level. In Fig. 3b we show the behaviour of T_{K-T} close to the QCP as a function of electron density. Quantum critical scaling¹⁴ predicts $T_{K-T} \propto (n - n_c)^{z\nu}$, where z is the dynamical exponent and ν is the correlation length exponent. As shown in Fig. 3b and Extended Data Fig. 3b and c, we find $z\nu \approx 2/3$. The same scaling was observed in other systems^{15–17}, and indicates that the superconductivity transition is of the classical three-dimensional XY or, equivalently, 2D quantum universality class^{16,17}. In the absence of a specific screening or dissipation mechanism^{18,19}, z is expected to be 1, so that $\nu = 2/3$, which implies that our system is in the clean limit by the Harris criterion²⁰.

The temperature dependence of the sheet resistance can be written as $R_S = R_{S0} + CT^\alpha$ (where R_{S0} is the residual resistance at 3 K), which decreases monotonically with charge-carrier density as shown in Fig. 3c, in accordance with the above conclusion regarding the Harris criterion. In an ordinary metal (or Fermi liquid) we expect $\alpha = 2$, independent of doping. Nevertheless, in Fig. 3c we find $1 < \alpha < 2$ (Extended Data Fig. 4a, b) over the entire phase diagram. Notice that at around $n = 7.5 \times 10^{14} \text{ cm}^{-2}$, $\alpha \cong 1.5$ extends down to temperatures close to the superconductivity transition, which would seem to indicate the presence of another QCP, owing to suppression of the CDW, inside the superconductivity dome. In what follows we show that this is not the case.

top-gate operation at 285 K. Sheet resistance R_S (in units of kΩ per square) is shown on the left and electron density δn_{2D} , measured by the Hall effect at 285 K, on the right. **c**, **d**, At high and low temperatures, respectively, the CDW and superconductivity transitions can be clearly identified.

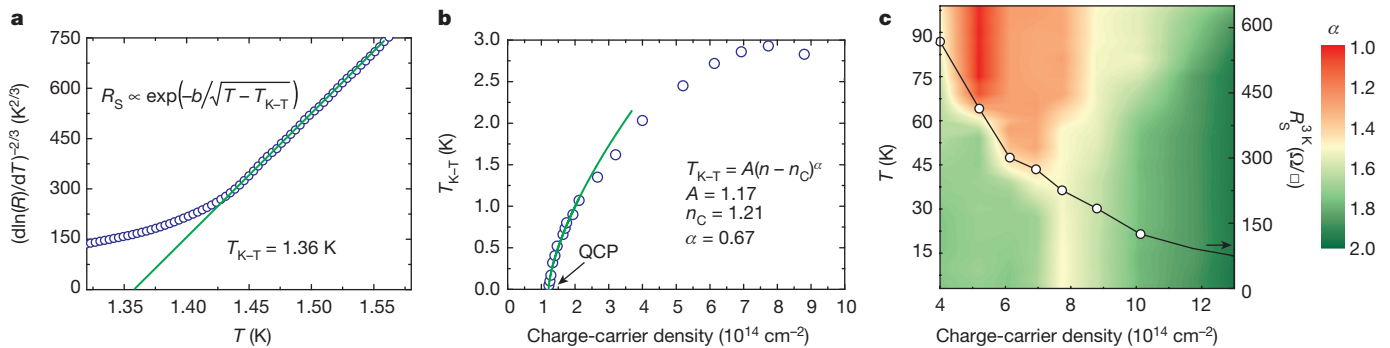


Figure 3 | Temperature dependence of the sheet resistance R_S close to the K-T transition. **a**, Fitting of the resistance to the K-T formula for a charge-carrier density of $n = 2.67 \times 10^{14} \text{ cm}^{-2}$. **b**, Behaviour of T_{K-T} at the QCP with critical exponent $\alpha \approx 2/3$. **c**, α values derived by fitting the

R versus T data from 3 K to 100 K with the formula $R_S = R_{S0} + CT^\alpha$, where α is plotted in colour against both charge-carrier density and temperature. The residual resistance at 3 K versus charge-carrier density is plotted on the right axis.

The presence of the competing orders in this 2D system has striking consequences for the electronic transport. In Fig. 4a we show the magnetoresistance as a function of magnetic field for a density of $n = 5.9 \times 10^{14} \text{ cm}^{-2}$. The magnetoresistance in the superconductivity phase is positive, as expected, but we clearly observe the presence of plateaus and oscillations in the data. By taking the derivative of the magnetoresistance, dR/dB , in Fig. 4b we observe that these features are temperature-independent and have well defined periods. The periodicity in magnetic field reflects a spatial periodicity given by the cyclotron equation, $\ell_c = (\Phi_0/\delta B)^{1/2}$, where $\Phi_0 = h/(2e) \approx 2,068 \text{ T nm}^2$ is the flux quantum and δB is the magnetic field periodicity. We have analysed the magnetoresistance data as a function of electron density (or gate voltage). (Magnetoresistance data for other electron densities are shown in Extended Data Fig. 5.) One can see a clear trend in the data (as shown in Fig. 4c): the length scale decreases monotonically with electron density from $\ell(n) \approx 450 \text{ nm}$ at $n \approx 1.3 \times 10^{14} \text{ cm}^{-2}$ to $\ell(n) \approx 170 \text{ nm}$ at

$n \approx 5.9 \times 10^{14} \text{ cm}^{-2}$ ($\ell(n)$ and δB are displayed in Fig. 4c as D_{CDW} and B_M , respectively).

It is clear that the well defined structure of the magnetoresistance reflects spatial fluctuations of the electronic pairing. A consistent explanation for these features is based on the Little-Parks effect²¹, whereby Cooper pairs are constrained to move in loops in the material—that is, pairing is local and constrained to well defined regions. The length scale we observe is associated with the trapping of magnetic flux quanta by the Cooper pairs. The existence of such a superconductivity matrix over a range of temperatures and charge-carrier densities in a single crystal is remarkable, leading us to suggest that the superconductivity matrix must be pinned and stabilized by an underlying matrix of inhomogeneous electronic states. Fluctuations of an underlying charge or spin order parameter have led to the discovery of superconductivity in a wide range of systems; the suppression of the CDW transition from 170 K to 40 K, concomitant with the appearance of superconductivity

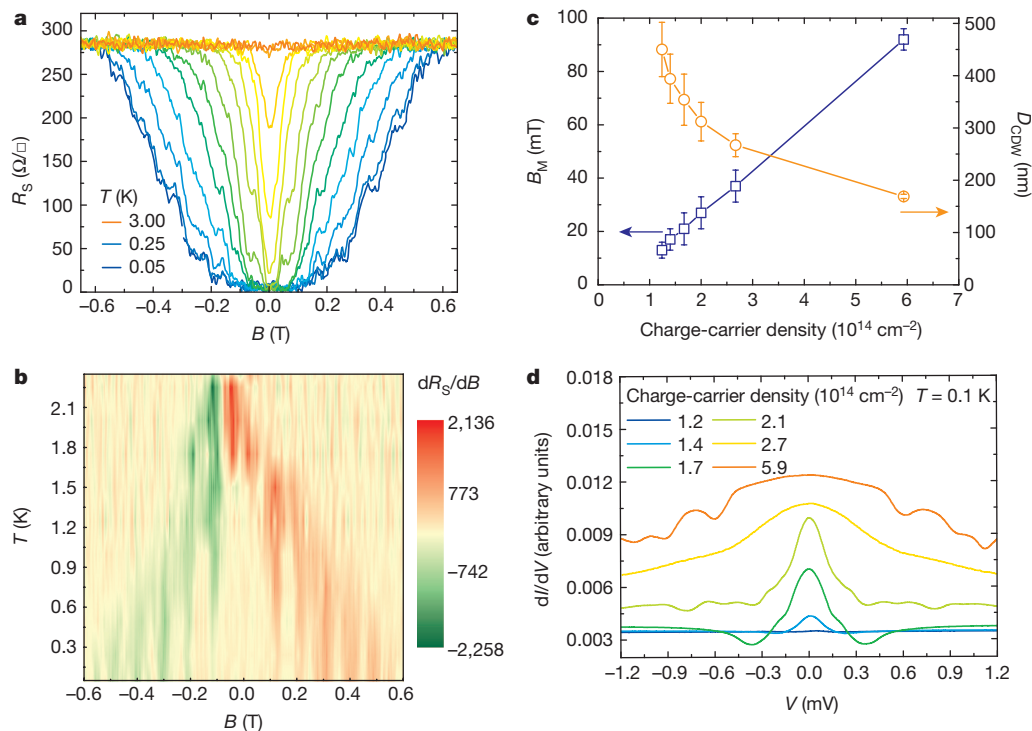


Figure 4 | The magnetoresistance for a charge-carrier density of $5.9 \times 10^{14} \text{ cm}^{-2}$ shows periodic oscillation. **a**, The magnetoresistance at different temperatures (each step is 0.25 K) in the superconductivity. **b**, Colour contour plot of the derivative of magnetoresistance dR_S/dB versus magnetic field B and temperature T , which displays temperature-independent periodic oscillations. **c**, Derived magnetoresistance

oscillating period B_M and the corresponding length scale D_{CDW} . Error bars define the 90% confidence interval. **d**, Charge-carrier-density-dependent two-terminal conductance dI/dV shows the ZBCP, indicating non- s -wave superconducting pair symmetry. Charge-carrier density $n = C \times 10^{14} \text{ cm}^{-2}$, where C is indicated by the colour; see legend.

and non-Fermi-liquid behaviour, strongly indicates that the CDW plays a part.

We now consider how local variations of the CDW can stabilize the superconductivity matrix. The CDW corresponds to a spatial modulation of the charge density $\delta\rho = \Delta(\mathbf{r})e^{-i[\mathbf{Q}(\mathbf{r}) \cdot \mathbf{r}]}$, where Δ is the CDW order parameter, \mathbf{r} is the position in the 2D plane and \mathbf{Q} is the CDW ordering vector. On the basis of symmetry alone, the Ginzburg–Landau free energy for Δ can be written as²²

$$F = \int d\mathbf{r} \left\{ a(\Delta)^2 + b(\Delta)^3 + c(\Delta)^4 + \frac{1}{2m^*Q^2} [|\mathbf{Q} \cdot (\nabla - i\mathbf{Q})\Delta|^2 + \kappa |\mathbf{Q} \times \nabla \Delta|^2] \right\} \quad (2)$$

where a, b, c, m^* and κ are phenomenological parameters that determine the energy scale for the spatial variation of ρ . Variations in the amplitude Δ are energetically very costly because the CDW has to be locally destroyed. However, variations in the CDW phase are energetically allowed and can be expressed as

$$\Delta(\mathbf{r}) = \Delta_0 \exp \left\{ i \left[\frac{\mathbf{K}}{2} \cdot \mathbf{r} - \theta(\mathbf{r}) \right] \right\} \quad (3)$$

where Δ_0 is the CDW order parameter in the uniform CCDW phase, \mathbf{K} is a reciprocal lattice vector in the direction of \mathbf{Q} , $\theta(\mathbf{r})$ is a spatially varying phase and we have included the known CCDW wavevector for TiSe_2 (1/2, 1/2, 1/2). When $\theta = \pi n$ (where n is an integer) we again have a CCDW, whereas if $\theta(\mathbf{r}) = \mathbf{Q} \cdot \mathbf{r}$ we have an ICDW. For illustrative purposes, we assume that the variation in $\theta(x)$ is one-dimensional in nature and substitute equation (3) into equation (2) to obtain

$$\delta F = \int dx \frac{1}{2} \left\{ [\partial_x \theta(x) - 1]^2 - g[1 - \cos(2\theta(x))] \right\} \quad (4)$$

where g depends on the Ginzburg–Landau parameters and $x = |\mathbf{K}/2 - \mathbf{Q}|r$ is the dimensionless length scale. The last equation reflects how the free energy changes locally with the phase. Minimizing equation (4) with respect to θ we obtain

$$\frac{d^2\theta}{dx^2} = -2g \sin(2\theta) \quad (5)$$

the differential equation for the pendulum (x plays the part of ‘time’), which has periodic solutions with period $\ell_{\text{ICDW}} \approx \pi/(g^{1/2}|\mathbf{K}/2 - \mathbf{Q}|)$. These spatially periodic variations of the phase represent Néel-like domain walls of ICDW between CCDW regions with different phase where ℓ_{ICDW} is the thickness of the domain wall. McMillan described²³ how these defects allow a uniform ICDW with slowly varying phase to break apart into domains of CCDW separated by ICDW domain walls that have more rapidly varying phases. The domain wall density is π/Q to match the homogeneous ICDW state.

A full solution to this problem in 2D is lacking, but we speculate that domain walls form a periodic matrix illustrated schematically in Fig. 1; blue CCDW regions with constant phases are embedded in a periodic ICDW matrix. We note that a similar structure was observed by scanning transmission microscope measurements of the closely related 1T-TaS₂, in which the ICDW state exists at ambient conditions²⁴. The self-organizing principle is that repulsive interactions occur between domain walls owing to higher-order terms in the free energy²³. Therefore the ICDW domains will form a matrix, breaking the CCDW into domains with fixed area, as required by the Little–Parks effect.

As shown by ref. 23, ICDW dynamic phase fluctuations—that is, phonon modes of the ICDW (not the lattice)—can exist in these domain walls. It is conceivable that these ICDW phonons induce superconductivity pairing and localize Cooper pairs in one-dimensional regions of the 2D system. Another intriguing aspect of our results

is displayed in Fig. 4d, which shows the point-contact conductance spectra measured at each density, in which we observe a clear zero bias conductance peak (ZBCP) in the superconductivity state. Extended Data Fig. 6 shows its temperature and magnetic field dependence at a density of $n = 2.1 \times 10^{14} \text{ cm}^{-2}$. ZBCPs are observed in a wide range of unconventional superconductors and are understood to arise by Andreev reflection from a Cooper pairing potential having an internal phase shift of the superconductivity order parameter²⁵. These results are therefore in stark contrast to the experimentally determined single-gap *s*-wave superconductivity observed in the Cu-intercalated Cu_xTiSe_2 (ref. 26). It is unlikely that our 2D samples would develop a superconductivity order parameter that is qualitatively distinct from that of Cu_xTiSe_2 (for example, *d*-wave). The existence of the ZBCP therefore suggests that, together with the spatial modulation of the superconductivity amplitude (which is demonstrated by the Little–Parks effect), there may also be a modulation of the superconductivity phase, although the correspondence between the amplitude variation and the phase variation cannot be determined from our measurements.

The observed state in 1T-TiSe₂ bears some similarity to the pair density wave (PDW) superconducting CDW phases. However, further experiments are required to substantiate the PDW hypothesis. Although one-dimensional PDW states have attracted more attention within the context of the copper oxide superconductors²⁷, more general PDW states having phase and amplitude variations in 2D are expected to be possible²⁸.

The coexistence of CCDW and ICDW was first observed by recent X-ray measurements of TiSe₂ at pressures close to where the superconductivity phase was expected²⁹. ICDW domain walls with a periodicity along the *c* axis of $\sim 300 \text{ nm}$ were observed, similar to the length scale determined in this experiment. While the periodicity was most pronounced along the *c* axis, a weak in-plane signal of incommensurability was observed that might correspond to the electronic microstructure observed here in the superconducting order (Abbamonte, P., personal communication, 15 January 2015).

In summary, we studied samples of TiSe₂ a few nanometres in thickness and tuned the material through the CDW and superconductivity phases using the electric-field effect. This technique allowed us to study in great detail the QCP in the material and classify its universality class. We also identified the interplay between superconductivity and CDW through the formation of an inhomogeneous many-body state which we identify with the localization of Cooper pairs along a matrix of incommensurate dislocations surrounding regions of CCDW. We conjecture that the superconductivity has in its origin in the coupling of the McMillan phonon modes of the ICDW with the electrons. These results open up opportunities for electric-field tuning of many-body states in condensed matter research.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 April; accepted 19 October 2015.

Published online 23 December 2015; corrected online 13 January 2016

(see full-text HTML version for details).

1. Ueno, K. *et al.* Electric-field-induced superconductivity in an insulator. *Nature Mater.* **7**, 855–858 (2008).
2. Bollinger, A. T. *et al.* Superconductor–insulator transition in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ at the pair quantum resistance. *Nature* **472**, 458–460 (2011).
3. Yu, Y. *et al.* Gate-tunable phase transitions in thin flakes of 1T-TaS₂. *Nature Nanotechnol.* **10**, 270–276 (2014).
4. Wang, Q. H., Kalantar-Zadeh, K., Kis, A., Coleman, J. N. & Strano, M. S. Electronics and optoelectronics of two-dimensional transition metal dichalcogenides. *Nature Nanotechnol.* **7**, 699–712 (2012).
5. Shen, K. M. & Davis, J. C. S. Cuprate high- T_c superconductors. *Mater. Today* **11**, 14–21 (2008).
6. Stewart, G. R. Superconductivity in iron compounds. *Rev. Mod. Phys.* **83**, 1589–1652 (2011).
7. Stewart, G. R. Heavy-fermion systems. *Rev. Mod. Phys.* **56**, 755–787 (1984).
8. Castro Neto, A. H. & Jones, B. A. Non-Fermi-liquid behavior in U and Ce alloys: criticality, disorder, dissipation, and Griffiths-McCoy singularities. *Phys. Rev. B* **62**, 14975–15011 (2000).

9. Kotov, V. N., Uchoa, B., Pereira, V. M., Guinea, F. & Castro Neto, A. H. Electron-electron interactions in graphene: current status and perspectives. *Rev. Mod. Phys.* **84**, 1067–1125 (2012).
10. Fisher, M. E. & Langer, J. S. Resistive anomalies at magnetic critical points. *Phys. Rev. Lett.* **20**, 665–668 (1968).
11. Castro Neto, A. H. Charge density wave, superconductivity, and anomalous metallic behavior in 2D transition metal dichalcogenides. *Phys. Rev. Lett.* **86**, 4382–4385 (2001).
12. Minnhagen, P. The two-dimensional Coulomb gas, vortex unbinding, and superfluid-superconducting films. *Rev. Mod. Phys.* **59**, 1001–1066 (1987).
13. Reyren, N. *et al.* Superconducting interfaces between insulating oxides. *Science* **317**, 1196–1199 (2007).
14. Sachdev, S. *Quantum Phase Transitions* (Cambridge Univ. Press, 1999).
15. Caviglia, A. D. *et al.* Electric field control of the $\text{LaAlO}_3/\text{SrTiO}_3$ interface ground state. *Nature* **456**, 624–627 (2008).
16. Parendo, K. A. *et al.* Electrostatic tuning of the superconductor-insulator transition in two dimensions. *Phys. Rev. Lett.* **94**, 197004 (2005).
17. Aubin, H. *et al.* Magnetic-field-induced quantum superconductor-insulator transition in $\text{Nb}_{0.15}\text{Si}_{0.85}$. *Phys. Rev. B* **73**, 094521 (2006).
18. Mason, N. & Kapitulnik, A. Superconductor-insulator transition in a capacitively coupled dissipative environment. *Phys. Rev. B* **65**, 220505 (2002).
19. Yazdani, A. & Kapitulnik, A. Superconducting-insulating transition in two-dimensional a -MoGe thin films. *Phys. Rev. Lett.* **74**, 3037–3040 (1995).
20. Harris, A. B. Effect of random defects on the critical behaviour of Ising models. *J. Phys. Chem.* **7**, 1671 (1974).
21. Tinkham, M. *Introduction to Superconductivity* (McGraw-Hill, 1975).
22. McMillan, W. L. Theory of discommensurations and the commensurate-incommensurate charge-density-wave phase transition. *Phys. Rev. B* **14**, 1496–1502 (1976).
23. McMillan, W. L. Time-dependent Landau theory of charge-density waves in transition-metal dichalcogenides. *Phys. Rev. B* **12**, 1197–1199 (1975).
24. Burk, B., Thomson, R. E., Zettl, A. & Clarke, J. Charge-density-wave domains in 1T-TaS₂ observed by satellite structure in scanning-tunneling-microscopy images. *Phys. Rev. Lett.* **66**, 3040 (1991).
25. Kashiwaya, S. & Tanaka, Y. Tunnelling effects on surface bound states in unconventional superconductors. *Rep. Prog. Phys.* **63**, 1641–1724 (2000).
26. Li, S. Y., Wu, G., Chen, X. H. & Taillefer, L. Single-gap s-wave superconductivity near the charge-density-wave quantum critical point in Cu_xTiSe_2 . *Phys. Rev. Lett.* **99**, 107001 (2007).
27. Fradkin, E., Kivelson, S. A. & Tranquada, J. M. Theory of intertwined orders in high temperature superconductors. *Rev. Mod. Phys.* **87**, 457 (2015).
28. Agterberg, D. F. & Tsunetsugu, H. Dislocations and vortices in pair-density-wave superconductors. *Nature Phys.* **4**, 639–642 (2008).
29. Joe, Y. I. *et al.* Emergence of charge density wave domain walls above the superconducting dome in 1T-TiSe₂. *Nature Phys.* **10**, 421–425 (2014).

Acknowledgements We thank L. Q. Chu, T. H. Ren, J. Y. Tan, J. Wu and H. Schmidt for experimental assistance, and S. Natarajan for assistance in preparing the manuscript. A.H.C.N. acknowledges many discussions with A. K. Geim and a private communication with P. Abbamonte. K.P.L. acknowledges a MOE Tier 1 grant, “2-D crystals as a platform for optoelectronics (R-143-000-556-112)”. G.E. acknowledges a National Research Foundation (NRF) Research Fellowship (NRF-NRFF2011-02). B.Ö. acknowledges support by the NRF, Prime Minister’s Office, Singapore, under its Competitive Research Programme (CRP award number NRF-CRP9-2011-3), and the SMF-NUS Research Horizons Award 2009-Phase II. A.H.C.N. acknowledges the CRP award, “Novel 2D materials with tailored properties: beyond graphene” (NRF-CRP6-2010-05). All authors acknowledge the NRF, Prime Minister’s Office, Singapore, under its Medium-Sized Centre Programme.

Author Contributions L.J.L. performed the growth and characterization of the single crystal, L.J.L. and E.C.T.O’F. performed device fabrication and carried out the measurement, L.J.L., E.C.T.O’F., K.P.L., B.Ö. and A.H.C.N. analysed the data and wrote the manuscript. All authors commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.H.C.N. (phycast@nus.edu.sg).

METHODS

Crystal growth and quality verification. TiSe₂ single crystals were grown in two steps by the chemical vapour transport method³⁰. First, polycrystalline TiSe₂ was prepared by mixing high-purity titanium powder (from Alfa Aesar, 99.99%) and selenium powder (from Alfa Aesar, 99.999%) in a stoichiometric ratio and heating the mixture at 800 °C for 3 days in a vacuum-sealed (<10⁻⁶ Torr) silica tube. Second, the polycrystalline powder was loaded into a two-zone tube furnace together with the transport agent I₂ at a concentration of 5 mg cm⁻³. The polycrystalline powder was then heated to 670 °C and single crystals of TiSe₂ were collected at 600 °C over a period of 10 days.

The quality of the bulk single crystals was confirmed by X-ray diffraction and temperature-dependent Raman spectroscopy³¹, as shown in Extended Data Fig. 1a and b. Further energy dispersive X-ray spectroscopy verified the stoichiometric composition of the crystals.

Device fabrication and characterization. TiSe₂ was exfoliated in a pure argon atmosphere by Scotch tape³² onto a SiO₂ (300 nm)/Si wafer and examined under high-resolution optical microscope. The non-uniformity in thickness can be discriminated by cross-correlation of the colour with atomic force microscope measurements of the height. Flakes with uniform thickness of around 10 nm or less and a long bar shape were selected for the device fabrication. Electrodes for transport measurements were fabricated by standard electron beam lithography techniques using a polymethylmethacrylate (PMMA) positive resist, followed by deposition of Ti (10 nm)/Au (65 nm). Thin crystals (one to three layers) of commercial hexagonal boron nitride were transferred onto the nanosheets within the argon atmosphere^{33,34}; the role of hexagonal boron nitride is to protect the TiSe₂ from degradation by both oxidation and damage by the electrolyte gate.

Atomic force microscope results show that the surface is clean (as shown in Extended Data Fig. 2b), with a roughness within ±1 nm, which may result from the non-uniform thickness of the TiSe₂ flake.

Electrical transport measurements were performed in both a ⁴He cryostat and in a ³He/⁴He dilution cryostat. Electrical transport measurements were performed using standard alternating-current (a.c.) lock-in amplifier and direct-current (d.c.) techniques, and resistance-versus-temperature and field measurements were performed using currents of 10–100 nA to avoid Joule heating.

The ion gel solution was prepared by mixing the triblock copolymer polystyrene-polymethylmethacrylate-polystyrene (PS-PMMA-PS) and the ionic liquid 1-ethyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide (EMIM-TFSI) into an ethyl propionate solvent (the weight ratio of polymer to ionic liquid to solvent is 0.7:9.3:90)³⁵. After covering the device with ion gel droplets by drop casting, the device as shown in Extended Data Fig. 2a was loaded into the cryostat and kept at room temperature and high vacuum for one hour to remove residual water from the electrolyte. Afterwards, resistance was measured against gate voltage to characterize the capability of the ion gel; a typical electrolyte gate sweep is shown in Fig. 1b.

The charge-carrier density doping by the ionic gate can be derived from the Hall-effect measurement both at high (285 K) and low (3 K) temperature; the former is shown in Extended Data Fig. 2c, and the latter is used to construct the phase diagram because the latter has a better direct correlation with the superconducting dome. Although the hexagonal boron nitride passivation prevents the accumulation of ions directly at the surface of TiSe₂, this was not found to reduce the capacity of the gate much, as demonstrated by our results and those of a recent work³⁶.

2D superconducting properties and the K–T transition. As discussed in the main text, the superconducting transition under different fixed perpendicular magnetic fields was measured. Extended Data Fig. 3b shows the magnetoresistance plot for a charge-carrier density of $n = 2.67 \times 10^{14} \text{ cm}^{-2}$. The upper critical field $H_{C2}(T)$ values can be determined from each curve at the intercept of extrapolations from the normal state and the superconductivity state. $H_{C2}(0)$ can be derived by interpolating the plot curve to zero temperature, which gives a value of 450 mT. The superconducting coherence length therefore can be derived from $\xi(0) = \sqrt{H_{C2}(0)/\phi_0}$. The minimum ξ that corresponds to the maximum H_{C2} point is about 22 nm, which is more than twice as large as the thickness of the measured device, indicating that superconductivity in our device is expected to have a 2D character.

For the K–T transition, the current–voltage response becomes nonlinear; a $V \propto I^\alpha$ relation is expected as a result of the vortex–antivortex pair unbinding³⁷. Consistent behaviour is also observed in our sample, as shown in Extended Data Fig. 3a.

To confirm the νz value that determines the nature of the quantum critical behaviour in our system, we also measured the temperature-dependent superconducting transition under perpendicular magnetic field at a fixed charge-carrier

density of $n = 2.67 \times 10^{14} \text{ cm}^{-2}$ (see Extended Data Fig. 3b). By using finite size scaling with the formula $R_S/R_C = F((B - B_C)T^{-1/\nu z})$, where R_C and B_C are two fitting parameters and F is an arbitrary function with $F(0) = 1$ (ref. 38), the data are expected to collapse into two sets of lines, with a certain νz value. As displayed in Extended Data Fig. 3c, the data collapse for $\nu z \approx 2/3$, which confirms the previous result.

Magnetoresistance oscillations at other doping levels. The magnetoresistance oscillation is observed when we sweep a perpendicular magnetic field at different temperature in the superconducting state. From the QCP point $n = 1.2 \times 10^{14} \text{ cm}^{-2}$ to the near-optimum doping $n = 5.9 \times 10^{14} \text{ cm}^{-2}$, the oscillations can be observed for all doping levels. However, these oscillations can only be clearly observed for certain temperatures T_0 and magnetic fields B_0 , whereas T_0 and B_0 values increase with increasing doping. For instance, for $n = 1.3 \times 10^{14} \text{ cm}^{-2}$, T_0 is 0.3 K and B_0 is 0.06 T; for $n = 2.7 \times 10^{14} \text{ cm}^{-2}$, T_0 increases to 0.4 K and B_0 increases to 0.13 T, as one can see from Extended Data Fig. 4. Although T_0 and B_0 values as well as the periods of oscillation δB increase with doping level, the amplitude of the magnetoresistance oscillation does not monotonically depend on doping levels. We find that the oscillating amplitude for doping levels of $1.3 \times 10^{14} \text{ cm}^{-2}$ and $5.9 \times 10^{14} \text{ cm}^{-2}$ is larger than that for other doping levels we measured. One can clearly see more contrast or sharpness for the periodic straight lines in Fig. 4b and Extended Data Fig. 5c than in Extended Data Fig. 5d. The stronger magnetoresistance oscillations at these doping levels could be related to the enhanced Cooper-pair phonon interaction, aroused by strong quantum fluctuation.

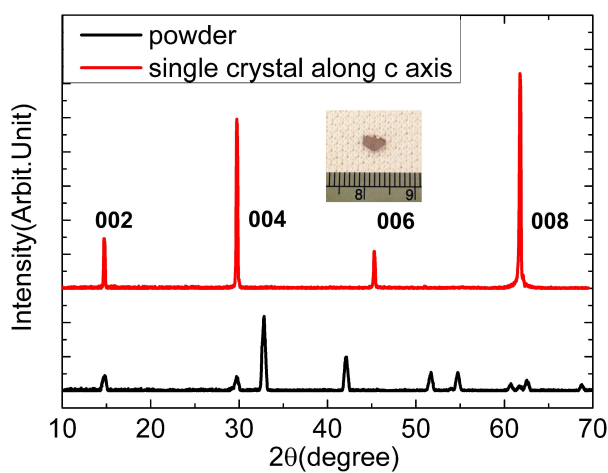
Temperature dependence of the sheet resistance. We plot the temperature dependence of the sheet resistance between 3 K and 100 K with the doping level ranging from $4 \times 10^{14} \text{ cm}^{-2}$ to $13 \times 10^{14} \text{ cm}^{-2}$ as shown in Extended Data Fig. 5a and b. By taking the temperature derivative $d(\log(R - R_0))/d(\log(T))$, α is extracted at each doping as a function of the temperature.

At doping levels away from optimal doping, $7.5 \times 10^{14} \text{ cm}^{-2}$, we observe Fermi-liquid behaviour at low temperatures below T_{CDW} . At the optimal doping level an exponent of 3/2 is observed over a wide range of temperature; this exponent is similar to the one observed in MnSi³⁹. As described in the main text, microscopic fluctuations of the order parameters from those of a CCDW to those of an ICDW gives rise to this temperature dependence.

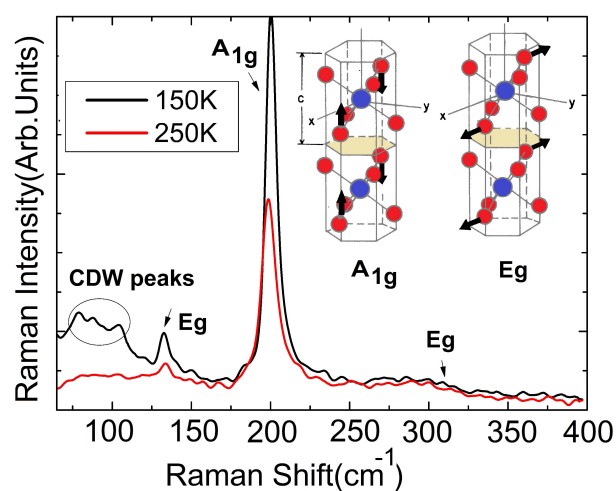
Point-contact conductance spectroscopy. Point-contact conductance spectroscopy of the normal-superconducting junction between Au/Ti and TiSe₂ was performed by the two-terminal a.c. + d.c. method, whereby the d.c. voltage is modulated with an additional a.c. voltage, such that the derivative dI/dV can be measured at the first harmonic by a current preamplifier and standard lock-in amplifier techniques.

The contacts were patterned by standard electron beam lithography using a PMMA positive resist. The development of the resist is performed in air to allow the oxidation of the contact region such that the contacts (despite not being nanoscale) are in the so-called ‘soft’ contact regime that has been successfully applied to pnictide and copper oxide superconductors⁴⁰. In this regime spectroscopic information can be obtained because the transport is primarily through multiple point-like pinholes whose individual dimension is smaller than the mean free path in the contact.

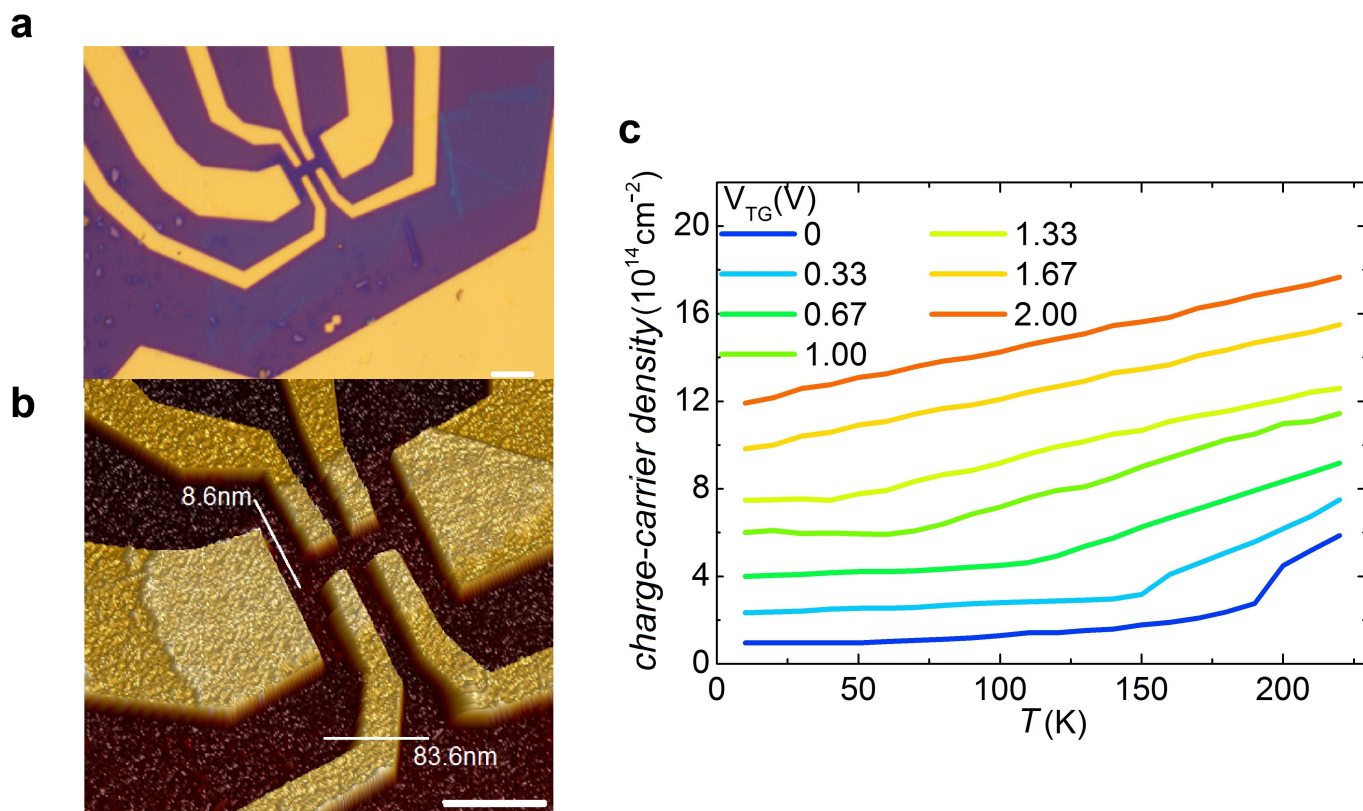
- Di Salvo, F. J., Moncton, D. E. & Waszczak, J. V. Electronic properties and superlattice formation in the semimetal TiSe₂. *Phys. Rev. B* **14**, 4321–4328 (1976).
- Holy, J. A., Woo, K. C., Klein, M. V. & Brown, F. C. Raman and infrared studies of superlattice formation in TiSe₂. *Phys. Rev. B* **16**, 3628–3637 (1977).
- Novoselov, K. S. *et al.* Electric field effect in atomically thin carbon films. *Science* **306**, 666–669 (2004).
- Doganov, R. A. *et al.* Transport properties of pristine few-layer black phosphorus by van der Waals passivation in an inert atmosphere. *Nature Commun.* **6**, 6647 (2015).
- Mayorov, A. S. *et al.* Micrometer-scale ballistic transport in encapsulated graphene at room temperature. *Nano Lett.* **11**, 2396–2399 (2011).
- Pu, J. *et al.* Highly flexible MoS₂ thin-film transistors with ion gel dielectrics. *Nano Lett.* **12**, 4013–4017 (2012).
- Gallagher, P. *et al.* A high-mobility electronic system at an electrolyte-gated oxide surface. *Nature Commun.* **6**, 6437 (2015).
- Reyren, N. *et al.* Superconducting interfaces between insulating oxides. *Science* **317**, 1196–1199 (2007).
- Fisher, M. P. A. Quantum phase transitions in disordered two-dimensional superconductors. *Phys. Rev. Lett.* **65**, 923–926 (1990).
- Pfleiderer, C., Julian, S. R. & Lonzarich, G. G. Non-Fermi-liquid nature of the normal state of itinerant-electron ferromagnets. *Nature* **414**, 427–430 (2001).
- Daghero, D. & Gonnelli, R. S. Probing multiband superconductivity by point-contact spectroscopy. *Supercond. Sci. Technol.* **23**, 043001 (2010).

a

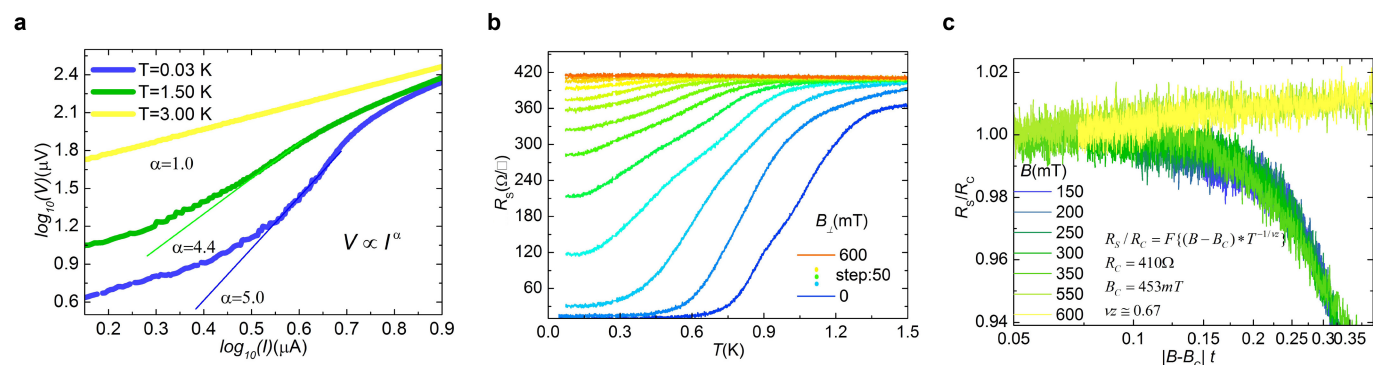
Extended Data Figure 1 | Characterization of the high quality of single-crystal TiSe₂. **a**, X-ray diffraction of both single-crystal and powder TiSe₂ sample. The inset shows the as-grown TiSe₂ single crystal. **b**, Raman spectroscopy pattern at both high temperature and low temperature.

b

The two main phonon modes, E_g and A_{1g}, are distinct, whereas only below T_{CDW} are the peaks corresponding to CDW phonon mode detectable. The inset displays the unit cell of the TiSe₂ lattice and the main phonon mode vectors.



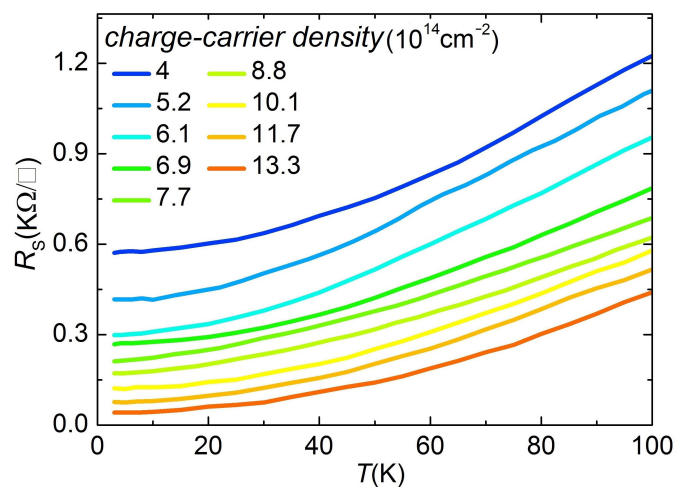
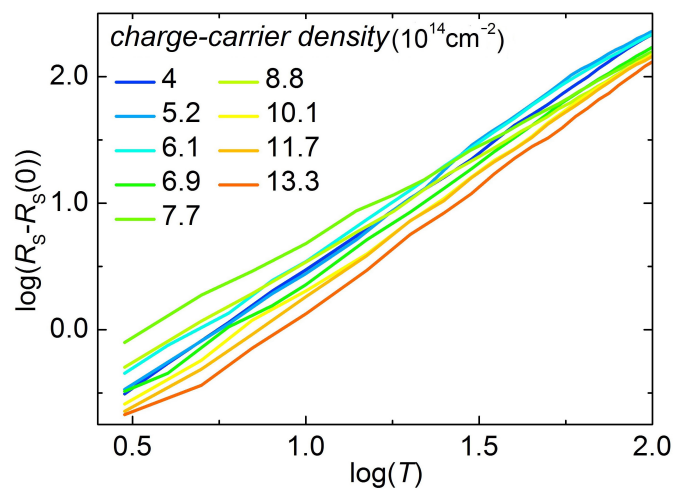
Extended Data Figure 2 | The Hall bar device and its characterization by Hall effect measurement. **a**, Optical microscope picture. **b**, Atomic force microscope picture of the Hall bar device. **c**, Temperature dependence of the charge-carrier density measured by the Hall effect at different top gate voltages, V_{TG} . Scale bar, 5 μm .



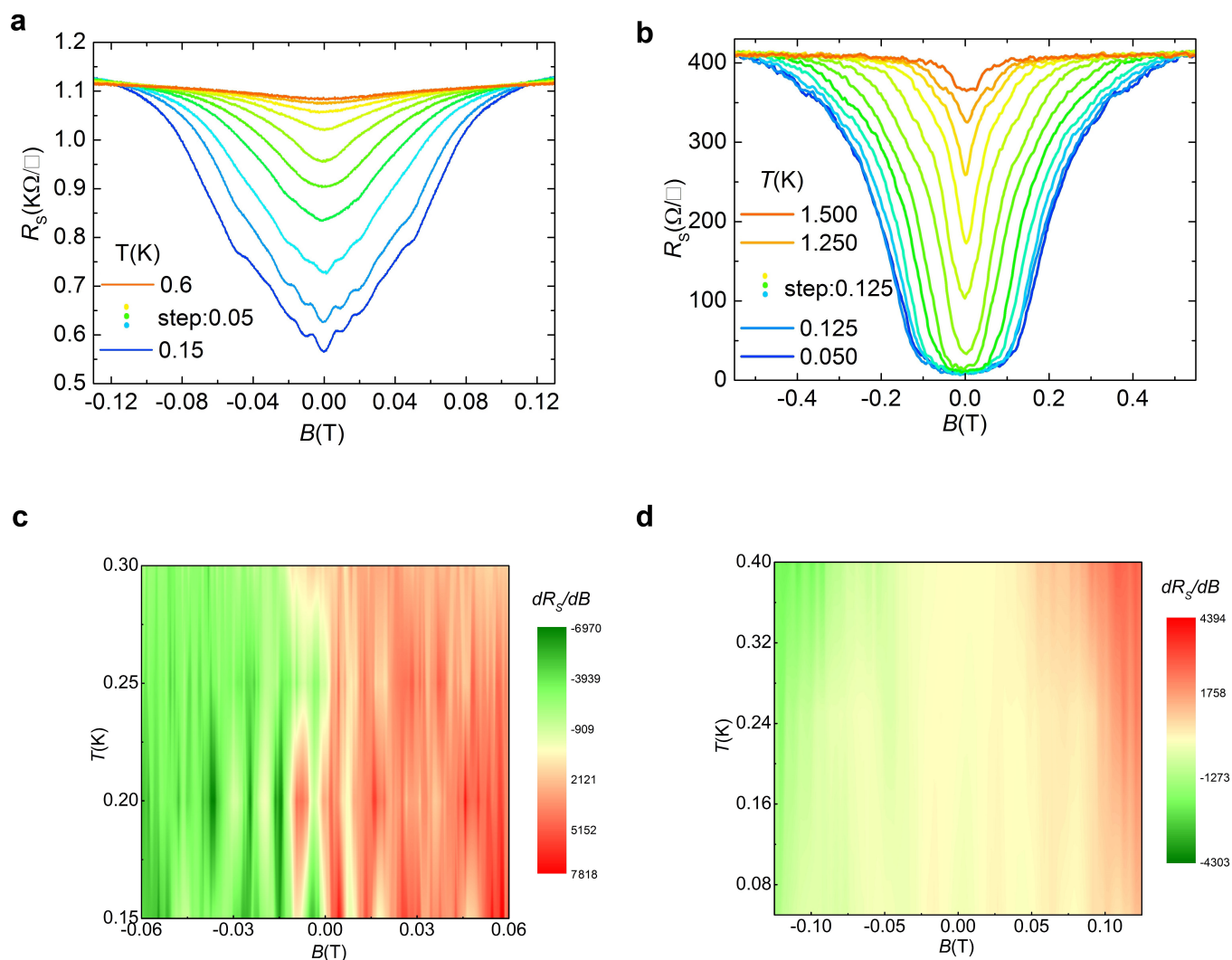
Extended Data Figure 3 | Characterization of the K-T transition.

- a**, The current–voltage power-law fit for $n = 5.9 \times 10^{14} \text{ cm}^{-2}$ at different temperatures is consistent with the behaviour of the 2D K–T transition.
- b**, Temperature-dependent magnetoresistance of the superconducting

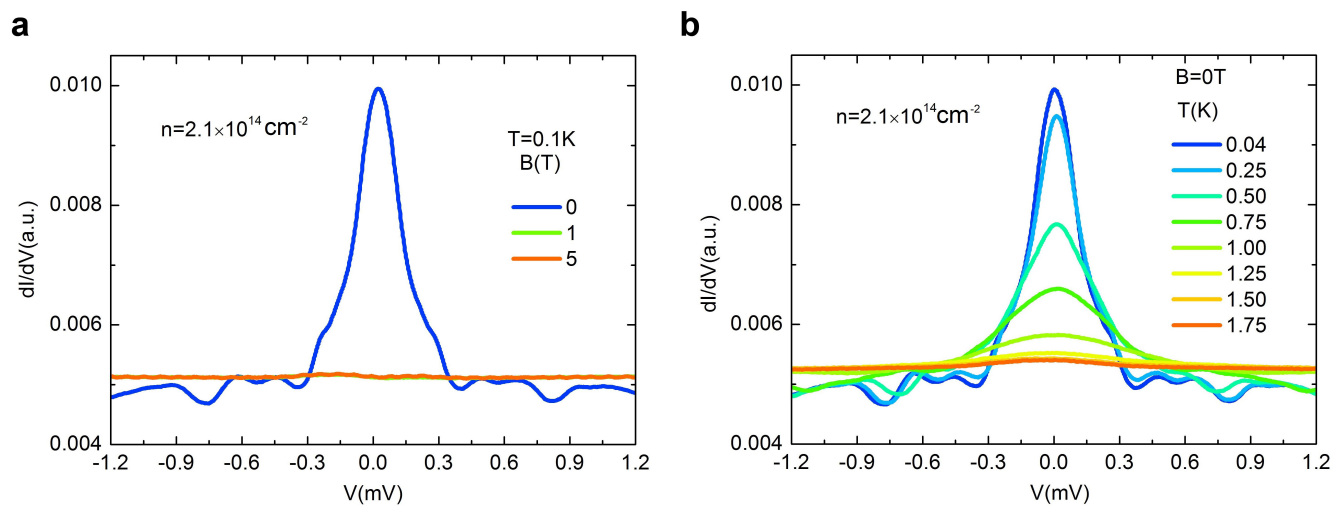
transition at different fixed perpendicular magnetic fields for $n = 2.67 \times 10^{14} \text{ cm}^{-2}$. **c**, The magnetoresistance data in **b** collapses into two sets of lines by so-called finite size scaling.

a**b**

Extended Data Figure 4 | The R versus T power-law fit indicates the existence of strong quantum fluctuation. a, Temperature dependence of the sheet resistance for different doping levels. **b,** The data shown in **a** is plotted on a log-log scale.



Extended Data Figure 5 | The magnetoresistance oscillation for charge-carrier densities of 1.3×10^{14} and $2.7 \times 10^{14} \text{ cm}^{-2}$. **a, c**, Perpendicular magnetic-field-dependent magnetoresistance measured at different temperatures. **b, d**, Plots of dR_s/dB against B and T for $n = 1.3 \times 10^{14} \text{ cm}^{-2}$ and $n = 2.7 \times 10^{14} \text{ cm}^{-2}$, respectively.



Extended Data Figure 6 | The conductance measured for a charge-carrier density of $2.1 \times 10^{14} \text{ cm}^{-2}$. **a**, Magnetic field dependence at 0.1 K. **b**, Temperature dependence at zero magnetic field. a.u., arbitrary units.

Rapid removal of organic micropollutants from water by a porous β -cyclodextrin polymer

Alaaeddin Alsaiee¹, Brian J. Smith¹, Leilei Xiao¹, Yuhan Ling², Damian E. Helbling² & William R. Dichtel¹

The global occurrence in water resources of organic micropollutants, such as pesticides and pharmaceuticals, has raised concerns about potential negative effects on aquatic ecosystems and human health^{1–5}. Activated carbons are the most widespread adsorbent materials used to remove organic pollutants from water but they have several deficiencies, including slow pollutant uptake (of the order of hours)^{6,7} and poor removal of many relatively hydrophilic micropollutants⁸. Furthermore, regenerating spent activated carbon is energy intensive (requiring heating to 500–900 degrees Celsius) and does not fully restore performance^{9,10}. Insoluble polymers of β -cyclodextrin, an inexpensive, sustainably produced macrocycle of glucose, are likewise of interest for removing micropollutants from water by means of adsorption¹¹. β -cyclodextrin is known to encapsulate pollutants to form well-defined host–guest complexes, but until now cross-linked β -cyclodextrin polymers have had low surface areas and poor removal performance compared to conventional activated carbons^{11–13}. Here we crosslink β -cyclodextrin with rigid aromatic groups, providing a high-surface-area, mesoporous polymer of β -cyclodextrin. It rapidly sequesters a variety of organic micropollutants with adsorption rate constants 15 to 200 times greater than those of activated carbons and non-porous β -cyclodextrin adsorbent materials^{7,8,11–13}. In addition, the polymer can be regenerated several times using a mild washing procedure with no loss in performance. Finally, the polymer outperformed a leading activated carbon for the rapid removal of a complex mixture of organic micropollutants at environmentally relevant concentrations. These findings demonstrate the promise of porous cyclodextrin-based polymers for rapid, flow-through water treatment.

Porous β -CD-containing polymers (P-CDPs) were derived from nucleophilic aromatic substitution of hydroxyl groups of β -CD by tetrafluoroterephthalonitrile (**1**). Although **1** has been copolymerized with bifunctional catechols previously¹⁴, its reaction with aliphatic alkoxides is undescribed. β -CD and **1** were polymerized in a suspension of K_2CO_3 in tetrahydrofuran (THF) at 80 °C to provide a pale-yellow precipitate in 20% yield, which proved to be a mesoporous high-surface-area polymer with the expected chemical bonds (Fig. 1a). The yield was further improved to 45% by performing the polymerization in THF:DMF (dimethylformamide) 9:1 by volume, in which β -CD is more soluble (see Supplementary Information section ‘Improved synthesis of P-CDP’ and Supplementary Figs 1 and 2). Following activation under high vacuum, N_2 porosimetry of the P-CDPs provided type II isotherms indicative of mesoporosity, and their Brunauer–Emmett–Teller surface areas (S_{BET}) ranged from 35 to 263 $m^2 g^{-1}$, depending on the molar feed ratio of 1: β -CD employed in the polymerization (Fig. 1b). P-CDPs obtained from a 1: β -CD ratio of 3:1 consistently exhibited the highest surface areas. Non-local density functional theory (NLDFT) calculations applied to the isotherms indicate that pores of 1.8–3.5 nm diameter comprise the majority of

the free volume of P-CDP (Fig. 1c), much like the pore size distributions of activated carbons (ACs; Extended Data Fig. 1). Alternative polymerization conditions (aqueous NaOH, 60 °C) produced a similar polymer that lacked permanent porosity (non-porous (NP)-CDP, $S_{BET} = 6 m^2 g^{-1}$), which serves as a useful control to demonstrate the importance of surface area for rapid micropollutant removal (see Methods section ‘Synthetic procedures’). ‘Water regain’ analysis of P-CDP and NP-CDP also reflected the higher pore volume of the former, which took up 265% of its weight when dispersed in H_2O as compared to 86% for the latter (Extended Data Table 1). Nevertheless, NP-CDP swells to a much greater degree, as its H_2O uptake is approximately 300 times its dry pore volume, as compared to a factor of only 23 for P-CDP. P-CDP’s combination of high H_2O uptake and modest swelling are desirable, as these parameters maximize adsorbent performance and minimize undesirable pressure drops associated with filtration processes.

Compositional analysis and spectroscopic characterization of the P-CDP and NP-CDP networks indicated the presence of both **1** and β -CD moieties in the polymers (see Methods section ‘FT-IR and solid-state ^{13}C NMR characterization of P-CDP and NP-CDP’, Extended Data Figs 2 and 3). The ratio of 1: β -CD in each polymer was determined by combustion analysis, with 6.1 equiv. of **1** per β -CD for P-CDP and 3.5 equiv. of **1** per β -CD for NP-CDP (see Methods section ‘Synthetic procedures’). Therefore P-CDP is more densely crosslinked than NP-CDP, which might be responsible for its permanent porosity. The F:N ratio also indicates that the terephthalonitrile moieties in each polymer are substituted by 2.1 and 2.2 alkoxides on average, which are consistent with model studies (see model reactions S1–S4 and Supplementary Figs 3–7), which suggest that the β -CD macrocycles are linked predominantly at both the small and large rims through disubstituted terephthalonitrile moieties.

The high surface area and permanent porosity of P-CDP enable the rapid removal of organic contaminants from water. Bisphenol A, a component of plastics that has attracted concerns as an endocrine disruptor, was chosen as a model pollutant to enable comparison with established adsorbents^{15,16}. We compared uptake of bisphenol A by P-CDP, NP-CDP, and a non-porous β -CD polymer crosslinked with epichlorohydrin (EPI-CDP, $S_{BET} = 23 m^2 g^{-1}$), which is the most extensively studied β -CD polymer for water purification and has been commercialized¹². We also tested three types of mesoporous ACs: the hybrid AC/ion exchange resin used in commercial Brita point-of-use filters (Brita AC, $S_{BET} = 507 m^2 g^{-1}$), DARCO granular activated carbon (GAC, $S_{BET} = 612 m^2 g^{-1}$), and Norit RO 0.8 activated carbon (NAC, $S_{BET} = 984 m^2 g^{-1}$), which is a leading AC typically used for high-value water purification (Extended Data Fig. 1). Each adsorbent (1 $mg ml^{-1}$) eventually removed most of the bisphenol A from a 0.1 mM (22.8 $mg l^{-1}$) aqueous solution, corresponding to equilibrium uptakes of 19–24 mg bisphenol A per g adsorbent (Extended Data Table 2, Extended Data Fig. 4), with P-CDP near the high end of this range (22 $mg g^{-1}$). More importantly, P-CDP removed bisphenol A

¹Department of Chemistry and Chemical Biology, Cornell University, Baker Laboratory, Ithaca, New York 14853, USA. ²School of Civil and Environmental Engineering, Cornell University, Ithaca, New York 14853, USA.

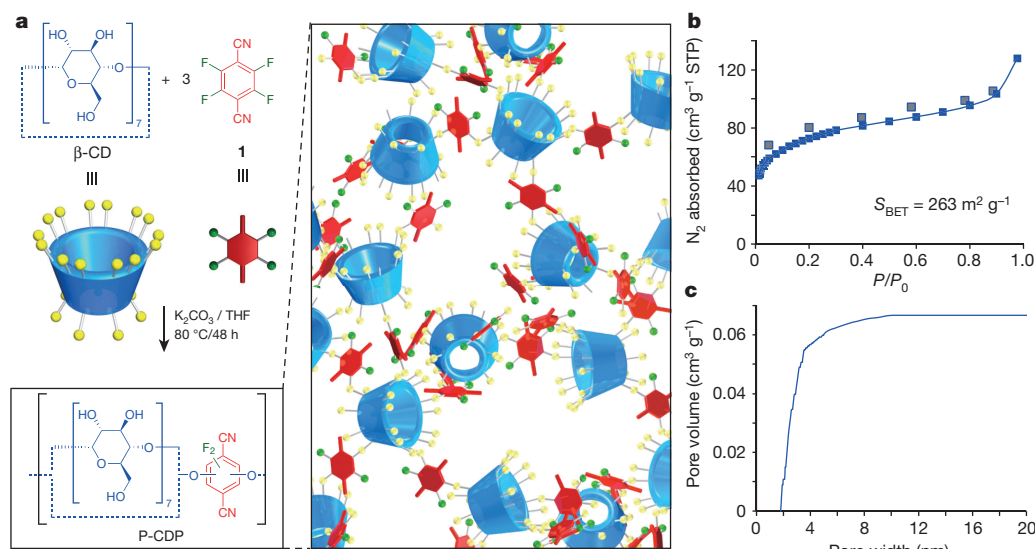


Figure 1 | β -CD polymer networks derived from nucleophilic aromatic substitution reactions. **a**, Left, synthesis of the high-surface-area porous P-CDP from β -CD and **1**. Right, schematic of the P-CDP structure. **b**, N_2 adsorption (blue squares) and desorption (grey squares) isotherms of P-CDP. The solid line is a guide to the eye. S_{BET} is the Brunauer–Emmett–

Teller (BET) surface area (in units of $m^2 g^{-1}$) of P-CDP calculated from the N_2 adsorption isotherm, and P and P_0 are the equilibrium and saturation pressures of N_2 at 77 K, respectively. **c**, The cumulative pore volume of P-CDP obtained by NLDFT analysis indicates the polymer's mesoporous structure.

much more quickly than all other adsorbents, reaching $\sim 95\%$ of its equilibrium uptake in 10 s (Fig. 2a). In contrast, NP-CDP required 30 min to reach equilibrium and adsorbed only 46% of its equilibrium value in 10 s, indicating that the near-instantaneous adsorption of bisphenol A by P-CDP is attributable to its porosity. Likewise, EPI-CDP required more than 1 h to reach equilibrium and only adsorbed 22% of its equilibrium value after 10 s, which is consistent with previous reports¹⁷. Finally, Brita AC and GAC each required more than 1 h to reach equilibrium, while NAC required 10 min (Fig. 2a). NAC only adsorbed 53% of its equilibrium value in 10 s despite its nearly four times higher surface area than P-CDP.

The apparent pseudo-second-order rate constant (k_{obs}) of bisphenol A adsorption to P-CDP is $1.5 mg g^{-1} min^{-1}$, which is 15 times higher than the high-performance NAC and two or more orders of magnitude higher than the other studied adsorbents (Extended Data Table 2, Extended Data Fig. 4). To our knowledge, this rate constant is the highest reported for bisphenol A or any other pollutant removed by ACs^{6,7}, mesoporous silicas¹⁸ or carbohydrate-based adsorbents^{19–21} under similar experimental conditions. k_{obs} amalgamates the

performance of readily accessible binding sites (conceptualized as the outer surface of the adsorbent) and less accessible binding sites (conceptualized as being within the adsorbent's interior). P-CDP's superior k_{obs} for bisphenol A adsorption indicates that nearly all of its β -CD binding sites are readily accessible, a feature not found in other adsorbents of which we are aware. We further probed the readily accessible binding sites of each adsorbent by determining the flow-through uptake of bisphenol A. In these experiments, the adsorbent ($\sim 3 mg$) was trapped as a thin layer on a $0.2 \mu m$ syringe filter, and aqueous bisphenol A (3 ml, 0.1 mM) was passed rapidly through the filter at a flow rate of $9 ml min^{-1}$. Under these conditions, P-CDP removed 80% of the bisphenol A from the solution, corresponding to more than 85% of its equilibrium uptake (Fig. 2b), whereas NAC removed 59% of the bisphenol A under the same conditions, indicative that nearly half of its binding sites are not accessible on the 20 s timescale. The superior performance of P-CDP further indicates that most of its β -CD moieties are rapidly accessed by bisphenol A.

The thermodynamic parameters of P-CDP's bisphenol A adsorption are consistent with the formation of β -CD inclusion complexes.

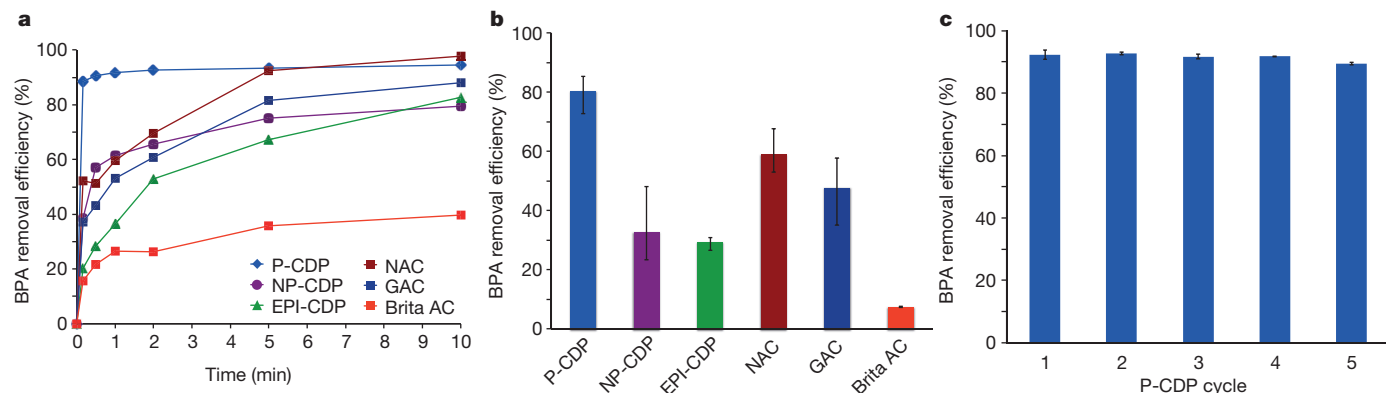


Figure 2 | Rate of bisphenol A uptake by various adsorbents. **a**, Time-dependent adsorption of aqueous bisphenol A (0.1 mM) by each adsorbent ($1 mg ml^{-1}$; see main text for adsorbent details). **b**, Removal of bisphenol A upon rapid flowing of the solution through a thin layer of the adsorbent. The data are reported as the average of triplicate experiments. Error bars, minimum and maximum removal.

c, The average percentage bisphenol A (BPA) removal efficiency by P-CDP after consecutive regeneration cycles. P-CDP was regenerated by rinsing the spent adsorbent with MeOH at room temperature. The data are reported as the average of triplicate experiments. Error bars, minimum and maximum removal.

P-CDP's equilibrium uptake as a function of residual bisphenol A concentration after adsorption, [BPA], fitted the Langmuir model (Methods section 'Thermodynamic studies of adsorption', Extended Data Fig. 5), suggesting 1:1 inclusion complex formation with an association constant (K) of $56,000 \text{ M}^{-1}$, which is comparable to the values reported for other β -CD polymers²². Furthermore, the maximum adsorption capacity at equilibrium ($q_{\text{max,e}}$) was found to be 88 mg g^{-1} , which is similar to the highest reported value of an EPI-CDP (84 mg g^{-1})¹¹, and corresponds to a bisphenol A: β -CD molar ratio of 0.9. Therefore, most of the β -CD units in the polymer are able to form 1:1 complexes with bisphenol A at equilibrium. At even higher concentrations of bisphenol A, P-CDP achieves bisphenol A: β -CD ratios greater than 1, presumably by binding bisphenol A on the outside of the CD rings or through other non-specific interactions. For example, 1 mg ml^{-1} of P-CDP adsorbs 200 mg g^{-1} of bisphenol A from a 1 mM aqueous solution, indicating significant capacity beyond 1:1 CD inclusion complexes. However, these results suggest that P-CDP's binding properties will reflect those of β -CD inclusion complex formation at concentrations relevant for water purification.

In contrast to the energy intensive and degradative regeneration processes of ACs, bisphenol A is easily removed from P-CDP by rinsing

the polymer with MeOH at room temperature. Five consecutive bisphenol A adsorption/desorption cycles were performed and the regenerated P-CDP exhibited almost no decrease in performance compared to the as-synthesized polymer (Fig. 2c). A functional cost analysis of P-CDP indicates raw materials costs of US dollar (USD) 3.70 per kg assuming further optimization of the polymerization, corresponding to manufacturing cost estimates of USD 5–25 per kg (see Supplementary Information section 'Functional cost analysis of P-CDP'). These estimates show promise compared to the costs (more than USD 9 per kg) of advanced ACs used for water treatment and wholesale costs of GAC (USD 22 per kg) and NAC (USD 47 per kg) used in this study. P-CDP's superior performance, facile regeneration procedure, and practical estimated cost make it reasonable to expect that it will prove economically competitive with ACs when full life-cycle analyses are performed.

In addition to bisphenol A, we evaluated the ability of P-CDP to remove pollutants of different size, functionality and hydrophobicity that span simple aromatics, pharmaceuticals and pesticides (Fig. 3a, Supplementary Table 1). The simple aromatics included the following: 2,4-dichlorophenol, an intermediate in herbicide production and degradation product of the antibacterial agent triclosan²³,

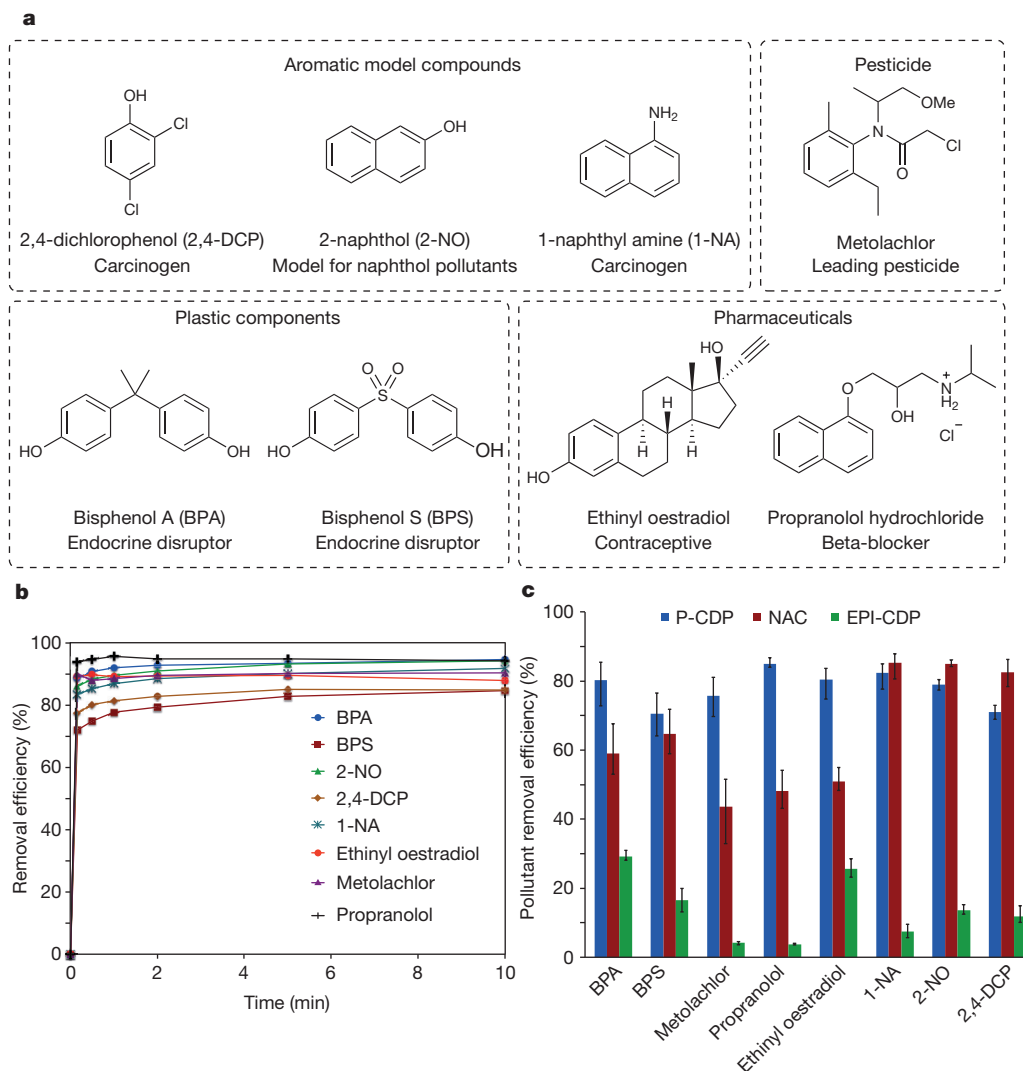


Figure 3 | Compound P-CDP rapidly adsorbs a broad range of organic micropollutants. **a**, Structures and relevance of each tested emerging organic micropollutant. Physicochemical properties most relevant to adsorption processes for each compound are available in Supplementary Table 2. **b**, Time-dependent adsorption of each pollutant (0.1 mM) by

P-CDP (1 mg ml^{-1}). **c**, Percentage removal efficiency of each pollutant obtained by rapidly flowing the adsorbate solution through a thin layer of P-CDP (blue), NAC (red) or EPI-CDP (green). The data are reported as the average uptake of triplicate experiments. Error bars, minimum and maximum uptake.

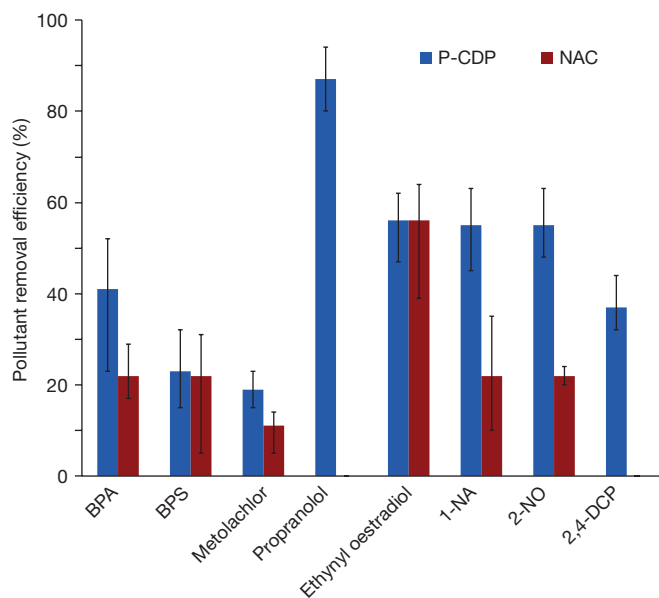


Figure 4 | P-CDP outperforms NAC for the rapid removal of a complex mixture of pollutants at environmentally relevant concentrations.

Percentage removal efficiency of a pollutant mixture obtained by rapidly flowing the adsorbate solution (8 ml) through a thin layer (0.3 mg) of P-CDP (blue) or NAC (red). Individual pollutant concentrations (in units of $\mu\text{g l}^{-1}$) in the mixture were: 100 (BPA); 2.5 (BPS); 5 (metolachlor); 100 (propranolol); 50 (ethinyl oestradiol); 5 (1-NA); 25 (2-NO); and 2.5 (2,4-DCP). Data are reported as the average uptake of three independent experiments. Error bars, minimum and maximum uptake.

1-naphthyl amine, an azo dye precursor and known carcinogen²⁴; and 2-naphthol, a model for various naphthol pollutants. We also evaluated the following anthropogenic contaminants: bisphenol S, which has replaced bisphenol A in many polycarbonates but also appears to be an endocrine disruptor with greater environmental persistence²⁵; metolachlor, one of the most common herbicides that is often detected in streams and groundwater²⁶; ethinyl oestradiol, an oestrogen mimic used in oral contraceptives that has caused the collapse of fish populations at concentrations as low as 5 ng l^{-1} (ref. 27); and propranolol, a beta-blocker used to treat hypertension, which is not removed efficiently by wastewater treatment protocols and has been detected in effluent streams at concentrations similar to blood serum levels of its users²⁸. Adsorption studies of each of these compounds were performed similarly to those for bisphenol A (0.1 mM adsorbate, 1 mg ml^{-1} adsorbent), except ethinyl oestradiol, which was tested at a lower concentration because of its low water solubility (0.04 mM adsorbate, 0.5 mg ml^{-1} adsorbent). Each organic contaminant is rapidly removed by P-CDP (Fig. 3b, Extended Data Fig. 6), and the time-dependent adsorption curves are similar to that of bisphenol A. The binding constants of the tested pollutants were estimated from the binding efficiency at equilibrium, and all were approximately 10^4 M^{-1} (Extended Data Tables 3 and 4). The rapid uptake of these pollutants by P-CDP was also investigated and compared with EPI-CDP and high performance NAC (Fig. 3c). P-CDP shows excellent rapid uptake of all pollutants, in stark contrast to the non-porous EPI-CDP, and it also outperforms NAC for all of the studied emerging contaminants. P-CDP even shows similar performance to NAC for the planar aromatic model compounds, which interact strongly with ACs. P-CDP's superior performance for anthropogenic contaminants indicates a major advantage of β -CD-based adsorbents: their three-dimensional cavities are a better match for non-planar compounds.

The rapid uptake of these pollutants was also investigated at environmentally relevant concentrations and in a mixture solution at concentrations between 2.5 and $100 \mu\text{g l}^{-1}$, the range in which many polar organic pollutants are quantified in wastewater²⁹ and drinking water

resources²⁷. The aqueous mixture of pollutants (8 ml) was passed rapidly through a $0.2 \mu\text{m}$ syringe filter containing approximately 0.3 mg of P-CDP or NAC (Fig. 4). On average, all of these emerging contaminants again showed equal or greater rapid uptake by P-CDP over NAC. Two pollutants showed no rapid uptake by NAC at low concentrations whereas all eight pollutants showed at least some removal by P-CDP. These results demonstrate that P-CDP can at least partially remove polar organic pollutants at environmental concentrations rapidly and simultaneously when present in mixtures, suggesting that it can contribute to the removal of a wide range of micropollutants during water and wastewater treatment.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 July; accepted 26 October 2015.

Published online 21 December 2015.

- Schwarzenbach, R. P. *et al.* The challenge of micropollutants in aquatic systems. *Science* **313**, 1072–1077 (2006).
- Richardson, S. D. & Ternes, T. A. Water analysis: emerging contaminants and current issues. *Anal. Chem.* **86**, 2813–2848 (2014).
- Murray, K. E., Thomas, S. M. & Bodour, A. A. Prioritizing research for trace pollutants and emerging contaminants in the freshwater environment. *Environ. Pollut.* **158**, 3462–3471 (2010).
- McKinlay, R., Plant, J. A., Bell, J. N. B. & Voulvoulis, N. Endocrine disrupting pesticides: implications for risk assessment. *Environ. Int.* **34**, 168–183 (2008).
- Daughton, C. G. & Ternes, T. A. Pharmaceuticals and personal care products in the environment: agents of subtle change? *Environ. Health Perspect.* **107**, 907–938 (1999).
- Órfão, J. J. M. *et al.* Adsorption of a reactive dye on chemically modified activated carbons — influence of pH. *J. Colloid Interf. Sci.* **296**, 480–489 (2006).
- Putra, E. K., Pranowo, R., Sunarso, J., Indraswati, N. & Ismadji, S. Performance of activated carbon and bentonite for adsorption of amoxicillin from wastewater: mechanisms, isotherms and kinetics. *Water Res.* **43**, 2419–2430 (2009).
- Kovalova, L., Knappe, D. R. U., Lehnberg, K., Kazner, C. & Hollender, J. Removal of highly polar micropollutants from wastewater by powdered activated carbon. *Environ. Sci. Pollut. Res.* **20**, 3607–3615 (2013).
- Chiang, P. C., Chang, E. E. & Wu, J. S. Comparison of chemical and thermal regeneration of aromatic compounds on exhausted activated carbon. *Water Sci. Technol.* **35**, 279–285 (1997).
- San Miguel, G., Lambert, S. D. & Graham, N. J. D. The regeneration of field-spent granular-activated carbons. *Water Res.* **35**, 2740–2748 (2001).
- Morin-Crini, N. & Crini, G. Environmental applications of water-insoluble β -cyclodextrin-epichlorohydrin polymers. *Prog. Polym. Sci.* **38**, 344–368 (2013).
- Lo Meo, P., Lazzara, G., Liotta, L., Riela, S. & Noto, R. Cyclodextrin-calixarene co-polymers as a new class of nanosponges. *Polym. Chem.* **5**, 4499–4510 (2014).
- Crini, G. & Morcellet, M. Synthesis and applications of adsorbents containing cyclodextrins. *J. Sep. Sci.* **25**, 789–813 (2002).
- Budd, P. M. *et al.* Polymers of intrinsic microporosity (PIMs): robust, solution-processable, organic nanoporous materials. *Chem. Commun.* 230–231 (2004).
- Vandenberg, L. N., Hauser, R., Marcus, M., Olea, N. & Welshons, W. V. Human exposure to bisphenol A (BPA). *Reprod. Toxicol.* **24**, 139–177 (2007).
- Liang, L. *et al.* Occurrence of bisphenol A in surface and drinking waters and its physicochemical removal technologies. *Front. Environ. Sci. Eng.* **9**, 16–38 (2015).
- Kitaoka, M. & Hayashi, K. Adsorption of bisphenol A by cross-linked β -cyclodextrin polymer. *J. Incl. Phenom. Macrocycl. Chem.* **44**, 429–431 (2002).
- Kim, Y.-H., Lee, B., Choo, K.-H. & Choi, S.-J. Selective adsorption of bisphenol A by organic-inorganic hybrid mesoporous silicas. *Micropor. Mesopor. Mater.* **138**, 184–190 (2011).
- Kyzas, G. Z., Lazaridis, N. K. & Bikiaris, D. N. Optimization of chitosan and β -cyclodextrin molecularly imprinted polymer synthesis for dye adsorption. *Carbohydr. Polym.* **91**, 198–208 (2013).
- Zhou, L.-C. *et al.* Highly efficient adsorption of chlorophenols onto chemically modified chitosan. *Appl. Surf. Sci.* **292**, 735–741 (2014).
- Wan Ngah, W. S., Teong, L. C. & Hanafiah, M. A. K. M. Adsorption of dyes and heavy metal ions by chitosan composites: a review. *Carbohydr. Polym.* **83**, 1446–1456 (2011).
- Aoki, N., Nishikawa, M. & Hattori, K. Synthesis of chitosan derivatives bearing cyclodextrin and adsorption of p-nonylphenol and bisphenol A. *Carbohydr. Polym.* **52**, 219–223 (2003).
- Latch, D. E. *et al.* Aqueous photochemistry of triclosan: formation of 2,4-dichlorophenol, 2,8-dichlorodibenzo-p-dioxin, and oligomerization products. *Environ. Toxicol. Chem.* **24**, 517–525 (2005).

24. Occupational Safety and Health Administration (OSHA) Standard, USA. *Toxic and Hazardous Substances: 13 Carcinogens (4-Nitrobiphenyl, etc.)*. Standard number 1910.1003. http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10007 (2012).
25. Ike, M., Chen, M. Y., Danzl, E., Sei, K. & Fujita, M. Biodegradation of a variety of bisphenols under aerobic and anaerobic conditions. *Water Sci. Technol.* **53**, 153–159 (2006).
26. Benner, J. *et al.* Is biological treatment a viable alternative for micropollutant removal in drinking water treatment processes? *Water Res.* **47**, 5955–5976 (2013).
27. Kidd, K. A. *et al.* Collapse of a fish population after exposure to a synthetic estrogen. *Proc. Natl Acad. Sci. USA* **104**, 8897–8901 (2007).
28. Kostich, M. S., Batt, A. L. & Lazorchak, J. M. Concentrations of prioritized pharmaceuticals in effluents from 50 large wastewater treatment plants in the US and implications for risk estimation. *Environ. Pollut.* **184**, 354–359 (2014).
29. Oulton, R. L., Kohn, T. & Cwiertny, D. M. Pharmaceuticals and personal care products in effluent matrices: a survey of transformation and removal during wastewater treatment and implications for wastewater management. *J. Environ. Monit.* **12**, 1956–1978 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by the National Science Foundation (NSF) through the Center for Sustainable Polymers (CHE-1413862). This research made use of the Cornell Center for Materials Research User Facilities, which are supported by the NSF (DMR-1120296). We acknowledge I. Keresztes for help with NMR spectroscopy, and M. Matsumoto for the design of the schematic of the polymer in Fig. 1a.

Author Contributions A.A., B.J.S., and L.X., and W.R.D. designed, synthesized, and characterized the cyclodextrin polymers and their micropollutant uptake at high concentrations. Y.L. and D.E.H. designed and conducted experiments that quantified micropollutant uptake at low concentrations. All authors wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.R.D. (wdichtel@cornell.edu) or D.E.H. (damian.helbling@cornell.edu).

METHODS

A complete set of detailed experiments, spectral data and adsorption experiments are available in Supplementary Information.

Reagents. β -cyclodextrin (β -CD) (>97%) and tetrafluoroterephthalonitrile (1) (>99%) were purchased from Sigma Aldrich and used without further purification. Tetrahydrofuran (THF) was purified and dried in a custom-built activated alumina solvent purification system. Epichlorohydrin (>99%) was purchased from Sigma Aldrich and used as received. Aqueous solutions of pollutants were prepared using 18 M Ω deionized H₂O at neutral pH. Pollutant model compounds were obtained from commercial sources and used as received. Norit RO 0.8 activated carbon (NAC) pellets were purchased from Sigma Aldrich and ground into a fine powder before use. Brita AC was obtained from a Brita Advanced Faucet point-of-use water filter and was ground into fine powder before use. Granular activated carbon (GAC; DARCO 12–20 mesh) was purchased from Sigma Aldrich, and ground into fine powder before use.

Materials and instrumentation. Pollutant removal experiments were performed at 25 °C on a stirring hot plate with a 250 r.p.m. stirring rate. Aqueous suspensions of adsorption experiments were filtered in syringes equipped with Whatman 0.2 μ m inorganic membrane filters. Instant pollutant removal experiments were also performed in syringes equipped with Whatman 0.2 μ m inorganic membrane filters.

Ultraviolet–visible (UV–vis) spectroscopy was performed on a Cary 5000 Varian UV–vis spectrometer. UV–vis spectra were recorded at RT over the range 200–600 nm, corrected against an appropriate background spectrum, and normalized to zero absorbance at 600 nm.

Quantification of analytes from the uptake of pollutant mixtures at μ g l^{−1} concentrations was performed by mass spectrometry (HPLC–MS). The analytical method was adopted from one previously reported for ultratrace level screening of polar and semi-polar organic chemicals³⁰ and involved high-performance liquid chromatography (HPLC) coupled with a quadrupole-orbitrap mass spectrometer (MS) (QExactive, ThermoFisher Scientific) and on-line solid phase extraction (EQuan Max Plus, ThermoFisher Scientific). Samples were injected at 5 ml volumes and were loaded onto an XBridge (Waters) C-18 Intelligent Speed (2.1 mm \times 20 mm, particle size 5 μ m) trap column. Elution from the trap column and onto an XBridge (Waters) C-18 analytical column (2.1 mm \times 50 mm, particle size 3.5 μ m) was performed using a gradient pump delivering 200 μ l min^{−1} of a water and MeOH mobile phase, each containing 0.1 vol.% formic acid. The HPLC–MS was operated with electrospray ionization in positive and negative polarity modes. The MS acquired full-scan MS data within a mass-to-charge range of 100–1,000 for each sample followed by a data-dependent acquisition of product ion spectra (MS/MS). Analytes were quantified from external calibration standards based on the analyte responses by linear least-squares regression. Limits of quantification for each analyte were determined as the lowest point in the external calibration curve at which at least 8 scans were measured across a chromatographic peak and the most intense MS/MS product ion was still detected. Exact molecular masses, ionization behaviour, retention times, and limits of quantification used for the detection and quantification of each analyte are provided in Supplementary Table 2.

Infrared spectroscopy was performed on a Thermo Nicolet iS10 with a diamond ATR attachment. Solution-phase NMR experiments were performed on a Varian INOVA-400 using a standard ¹H[¹³C, ¹⁵N] Z-PFG probe with a 20 Hz sample spin rate. Solid-state NMR analyses were conducted on a Varian INOVA-400 spectrometer using an external Kalmus ¹H linear pulse amplifier blanked using a spare line. Samples were packed into 7 mm outside diameter silicon nitride rotors and inserted into a Varian HX magic angle spinning (MAS) probe.

Surface area measurements were conducted on a Micromeritics ASAP 2020 Accelerated Surface Area and Porosimetry Analyzer. Each sample (25–50 mg) was degassed at 90 °C for 24 h and then backfilled with N₂. N₂ isotherms were generated by incremental exposure to ultrahigh-purity nitrogen up to 1 atm in a liquid nitrogen (77 K) bath, and surface parameters were determined using BET adsorption models included in the instrument software (Micromeritics ASAP 2020 V4.00).

Synthetic procedures. *Porous β -cyclodextrin polymer (P-CDP).* A flame dried 20 ml scintillation vial equipped with a magnetic stir bar was charged with β -CD (0.200 g, 0.176 mmol), 1 (0.100 g, 0.500 mmol), and K₂CO₃ (0.300 g, 2.17 mmol). The vial was flushed with N₂ gas for 5 min, then dry THF (8 ml) was added and the vial was bubbled with N₂ for additional 2–3 min. The N₂ inlet was removed and the mixture was placed on a hot stirring plate (85 °C) and stirred at 500 r.p.m. for 2 d. The orange suspension was cooled and then filtered, and the residual K₂CO₃ was removed by washing the solid on the filter paper with 1 N HCl until CO₂ evolution stopped. The recovered light yellow solid was isolated and activated by soaking in H₂O (2 \times 10 ml) for 15 min, THF (2 \times 10 ml) for 30 min and CH₂Cl₂ (1 \times 15 ml) for 15 min. Finally, the solid was dried under high vacuum at 77 K in a liquid nitrogen bath for 10 min and then at RT for 2–3 days. P-CDP (0.055 g, 20% yield)

was obtained as a pale yellow powder and subsequently characterized. ¹³C-MAS SS-NMR (400 MHz): chemical shifts δ of 168.9 p.p.m., 157.2 p.p.m., 131.1 p.p.m., 103.9 p.p.m., 95.2 p.p.m. and 71.8 p.p.m. relative to an external reference (the deuterium signal of liquid CDCl₃). IR (solid, attenuated total reflectance, ATR) 3,368, 2,937, 2,243, 1,684, 1,625, 1,478, 1,376, 1,304, 1,270, 1,153, 1,030 cm^{−1}. Analysis. Calculated for (C₄₂H₇₀O₃₅)₁•(C₈F₂N₂)₆•(CH₂Cl₂)₂•(H₂O)₂: C, 47.95; H, 3.17; F, 9.97; N, 7.35. Found: C, 48.23; H, 2.99; F, 9.66; N, 7.37.

Non-porous β -cyclodextrin polymer (NP-CDP). β -CD (2.00 g, 1.76 mmol) and 1 (2.11 g, 10.6 mmol) were mixed vigorously in an aqueous NaOH solution (6.25 N, 2.00 ml) at 85 °C. The mixture solidified within 1 h, after which deionized H₂O was added and the suspension was filtered. The solid was washed by soaking in deionized H₂O (2 \times 150 ml) for 15 min, THF (3 \times 15 ml) for 30 min and CH₂Cl₂ (1 \times 15 ml) for 15 min. Finally, the solid was dried under high vacuum at RT for 2 days to give NP-CDP (0.746 g, 20.1% yield) as a yellow powder. ¹³C-MAS SS-NMR (400 MHz): δ 162.9, 143.3, 140.4, 135.1, 117.0, 99.0, 96.2, 94.1, 72.6 p.p.m. IR (solid, ATR) 3,327, 2,938, 2,239, 1,674, 1,610, 1,463, 1,370, 1,268, 1,150, 1,100, 1,030 cm^{−1}. Analysis. Calculated for (C₄₂H₆₃O₃₅)₁•(C₈F_{1.8}N₂)_{3.5}•(H₂O)₁₃: C, 43.88; H, 4.68; F, 6.25; N, 5.12. Found: C, 43.78; H, 4.51; F, 6.31; N, 5.11.

Epichlorohydrin β -cyclodextrin polymer (EPI-CDP). β -CD (0.300 g, 2.64 mmol) was dissolved in aqueous NaOH (6.25 N, 5.00 ml) at 60 °C. Epichlorohydrin (2.50 ml, 32.4 mmol) was added to this solution dropwise while stirring vigorously at 60 °C. The mixture turned into a yellow gel within 1 h, after which 10 ml of deionized H₂O was added, and the mixture was filtered on a Büchner funnel. The solid was washed by soaking in deionized H₂O (2 \times 150 ml) for 15 min, THF (3 \times 15 ml) for 30 min and CH₂Cl₂ (1 \times 15 ml) for 15 min. The solid was finally dried under high vacuum for 2 d at RT to give EPI-CDP (3.11 g, 62% yield) as a white powder. ¹³C-MAS SS-NMR (400 MHz): δ 100.1, 72.0 p.p.m. IR (solid, ATR) 3,387, 2,923, 2,900, 1,702, 1,360, 1,030 cm^{−1}. Analysis. Calculated for (C₄₂H₆₀O₃₅)₁•(C₃H₆O)₁₀•(H₂O)_{4.5}: C, 48.40; H, 7.28. Found: C, 48.23; H, 7.09.

FT-IR and solid-state ¹³C NMR characterization of P-CDP and NP-CDP. FTIR spectra of P-CDP and NP-CDP showed absorbances at 2,235 cm^{−1}, corresponding to the nitrile stretch, as well as 1,670 cm^{−1} and 1,463 cm^{−1}, corresponding to C–C aromatic stretches. C–F stretches, which resonate at 1,268 cm^{−1}, are present in the spectra of both polymers and appear weaker compared to the spectrum of 1, as expected for partial F substitution. Finally, the IR spectra of P-CDP and NP-CDP exhibited O–H stretches near 3,330 cm^{−1}, aliphatic C–H stretches around 2,930 cm^{−1}, and an intense C–O stretch at 1,030 cm^{−1}, which are spectral features of intact β -CD (Extended Data Fig. 2). Solid-state ¹³C NMR spectra of P-CDP and NP-CDP exhibited resonances associated with β -CD at δ = 72 and 100 p.p.m. (Extended Data Fig. 3). Peaks at δ = 95 and 140 ppm correspond to the newly formed alkoxy groups and aromatic carbons, respectively.

Water regain analysis. P-CDP or NP-CDP (100 mg) were dispersed in deionized H₂O (10 ml) for 1 h and then filtered using 11 μ m Whatman filter paper. The solids were collected and blotted using additional Whatman filter paper, and weighed. The water regain (expressed as weight per cent) of each polymer was determined from the average of two measurements using the following equation:

$$\text{Water regain} = \frac{w_w - w_d}{w_d} \times 100$$

where w_w (mg) and w_d (mg) are the masses of the wet and dry polymer, respectively.

Batch adsorption kinetic studies. Adsorption kinetic studies were performed in 20 ml scintillation vials equipped with magnetic stir bars. All studies were conducted at ambient temperature on a stirring hot plate adjusted to provide a 250 r.p.m. stirring rate.

In studies involving P-CDP and NP-CDP, the polymer (18 mg) was initially washed with H₂O for 2–3 min and then filtered on 11 μ m Whatman filter paper. The polymer was then transferred to a 20 ml scintillation vial and then a pollutant stock solution (18 ml) was added. The mixture was immediately stirred and 2 ml aliquots of the suspension were taken at certain intervals via syringe and filtered immediately by a Whatman 0.2 μ m inorganic membrane filter. The residual concentration of the pollutant in each sample was determined by UV–vis spectroscopy. In studies involving EPI-CDP, NAC, GAC and Brita AC, the adsorbent (6 mg) was added to a 20 ml scintillation vial and then a pollutant stock solution (6 ml) was added. The vial was stirred for a measured amount of time before the suspension was filtered using a Whatman 0.2 μ m inorganic membrane filter.

The concentrations of pollutants in stock solutions as well as in the filtrates were characterized by UV–vis spectroscopy, based on calibration with their measured molar extinction coefficients (ϵ in units of M^{−1} cm^{−1}), which were determined for bisphenol A (3,343 at λ_{max} = 276 nm), bisphenol S (20,700 at λ_{max} = 259 nm), 2-naphthol (4,639 at λ_{max} = 273 nm), 1-naphthyl amine (5,185 at λ_{max} = 305 nm), 2,4-dichlorophenol (2,255 at λ_{max} = 284 nm), and metolachlor (213 at λ_{max} = 15,330 nm) pollutants. ϵ values of ethinyl oestradiol (8,430 at

$\lambda_{\max} = 220$ nm) and propranolol hydrochloride (5,310 at $\lambda_{\max} = 290$ nm) were reported elsewhere^{31,32}.

The efficiency of pollutant removal (in %) by the sorbent was determined by the following equation:

$$\text{Pollutant removal efficiency} = \frac{C_0 - C_t}{C_0} \times 100$$

where C_0 (mmol l⁻¹) and C_t (mmol l⁻¹) are the initial and residual concentration of pollutant in the stock solution and filtrate, respectively.

The amount of pollutant bound to the sorbent was determined by the following equation:

$$q_t = \frac{(C_0 - C_t)M_w}{m}$$

where q_t (mg g⁻¹) is amount of pollutant adsorbed per g of sorbent at time t (min). C_0 (mmol l⁻¹) and C_t (mmol l⁻¹) are the initial and residual concentration of pollutant in the stock solution and filtrate, respectively. m (g) is the mass of sorbent used in the study. M_w (g mol⁻¹) is the molar mass of the pollutant.

The uptake rate of each adsorbent was best described by Ho and McKay's pseudo-second-order adsorption model³³, shown in the following equation in a common linearized form:

$$\frac{t}{q_t} = \frac{t}{q_e} + \frac{1}{k_{\text{obs}}q_e^2}$$

where q_t and q_e are the adsorbate uptakes (mg adsorbate per g polymer) at time t (min) and at equilibrium, respectively, and k_{obs} is an apparent second-order rate constant (g mg⁻¹ min⁻¹).

Flow-through adsorption experiments. *Individual pollutants at high (mM) concentrations.* 3.0 mg of the adsorbent was stirred in 3 ml deionized H₂O for 2–3 min, then the suspension was pushed by a syringe through a Whatman 0.2 μm inorganic membrane filter to form a thin layer of the adsorbent on the filter membrane. 3 ml of the pollutant stock solution was then pushed through the adsorbent over 20 s (8–9 ml min⁻¹ flow rate). The filtrate was then measured by UV–vis spectroscopy to determine the pollutant removal efficiency.

Mixture of pollutants at environmentally relevant (μg l⁻¹) concentrations. 15 mg of the adsorbent (P-CDP or NAC) was added into a 20 ml vial and 5 ml nanopure water was added to prepare 3 g l⁻¹ stock suspension. Then 0.1 ml of the suspension was pushed through a Whatman 0.2 μm inorganic membrane filter with a syringe to form a thin layer of the adsorbent on the membrane. 8 ml of the diluted mixture (100 μg l⁻¹ bisphenol A, 2.5 μg l⁻¹ BPS, 50 μg l⁻¹ ethinyl oestradiol, 100 μg l⁻¹ propranolol, 5 μg l⁻¹ metolachlor, 5 μg l⁻¹ 1-Na, 25 μg l⁻¹ 2-NO, and 2.5 μg l⁻¹ 2,4-DCP) was then pushed through the adsorbent over approximately 20 s (25 ml min⁻¹ flow rate). The experiments were conducted in duplicate/triplicate. The filtrate was then measured by HPLC-MS (Supplementary Table 2).

Thermodynamic studies of adsorption. 4.0 mg of sorbent was initially washed with 3 ml of deionized H₂O for 2–3 min and then filtered on a Whatman filter paper. Then the solid was transferred to a 4 ml vial equipped with a stirring bar, and 2 ml (for 2 mg ml⁻¹ studies) or 4 ml (for 1 mg ml⁻¹ studies) of pollutant stock solution was added, and the suspension was stirred for 10 min to reach equilibrium. The suspension was then filtered on a Whatman 0.2 μm inorganic membrane filter, and the filtrate was measured by UV–vis spectroscopy.

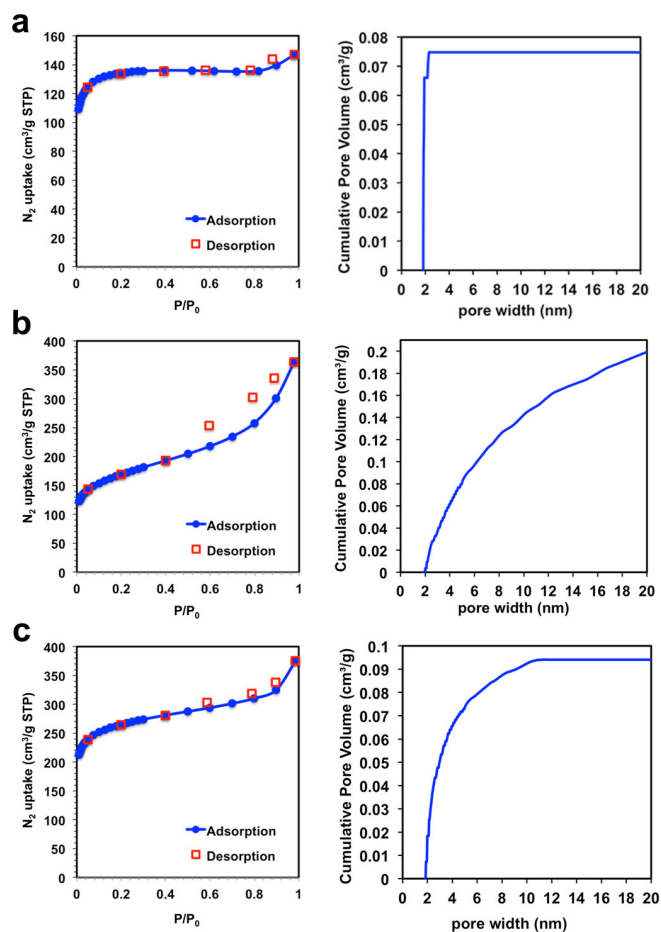
A Langmuir adsorption isotherm^{34,35} was generated by plotting $1/q_e$ versus $1/c$ in the following equation:

$$\frac{1}{q_e} = \frac{1}{q_{\text{max},e}} + \frac{1}{q_{\text{max},e}Kc}$$

where q_e (mg g⁻¹) is the amount of pollutant adsorbed at equilibrium, $q_{\text{max},e}$ (mg g⁻¹) is the maximum adsorption capacity of adsorbent at equilibrium, c (mol l⁻¹) is the residual pollutant concentration at equilibrium, and K (mol⁻¹) is the equilibrium constant.

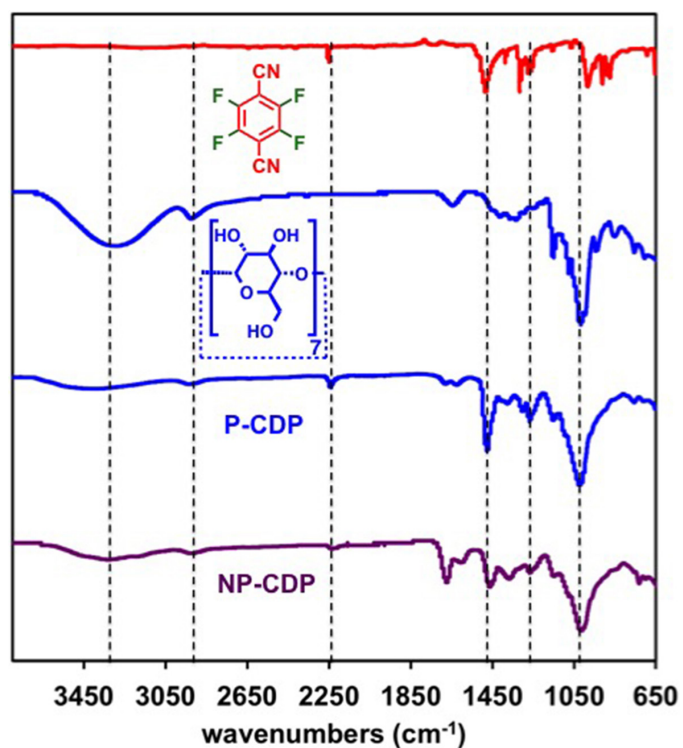
P-CDP regeneration experiments. 10 mg of P-CDP was initially soaked in 5 ml deionized H₂O for 5 min, and then filtered on a Whatman filter paper. The polymer was then transferred to a 20 ml scintillation vial equipped with a magnetic stir bar, to which a bisphenol A stock solution (10 ml, 0.1 mM) was added. The mixture was stirred at RT for 10 min, and then filtered on a Whatman filter paper. The residual bisphenol A concentration in the filtrate was measured by UV–vis. P-CDP was regenerated by soaking in MeOH (10 ml) for 5 min and recovered by filtration. This adsorption/desorption cycle was performed five times to generate the plot shown in Fig. 2c. The MeOH washing filtrate from the first cycle was concentrated under vacuum, and the residual solid was dissolved in 10 ml deionized H₂O and measured by UV–vis to determine the amount of recovered bisphenol A that was adsorbed on the polymer.

30. Huntscha, S., Singer, H. P., McArdell, C. S., Frank, C. E. & Hollender, J. Multiresidue analysis of 88 polar organic micropollutants in ground, surface and wastewater using online mixed-bed multilayer solid-phase extraction coupled to high performance liquid chromatography-tandem mass spectrometry. *J. Chromatogr. A* **1268**, 74–83 (2012).
31. Bisby, R., Botchway, S., Crisostomo, A., Parker, A. & Scherer, K. Fluorescence lifetime imaging of propranolol uptake in living glial C6 cells. *Spectrosc. Int. J.* **27**, 533–540 (2012).
32. O'Neil, M. J. (ed.) *The Merck Index — An Encyclopedia of Chemicals, Drugs, and Biologicals* (Merck and Co., 2006).
33. Ho, Y. S. & McKay, G. Pseudo-second order model for sorption processes. *Process Biochem.* **34**, 451–465 (1999).
34. Liu, H., Cai, X., Wang, Y. & Chen, J. Adsorption mechanism-based screening of cyclodextrin polymers for adsorption and separation of pesticides from water. *Water Res.* **45**, 3499–3511 (2011).
35. García-Zubiri, I. X., González-Gaitano, G. & Isasi, J. R. Sorption models in cyclodextrin polymers: Langmuir, Freundlich, and a dual-mode approach. *J. Colloid Interface Sci.* **337**, 11–18 (2009).

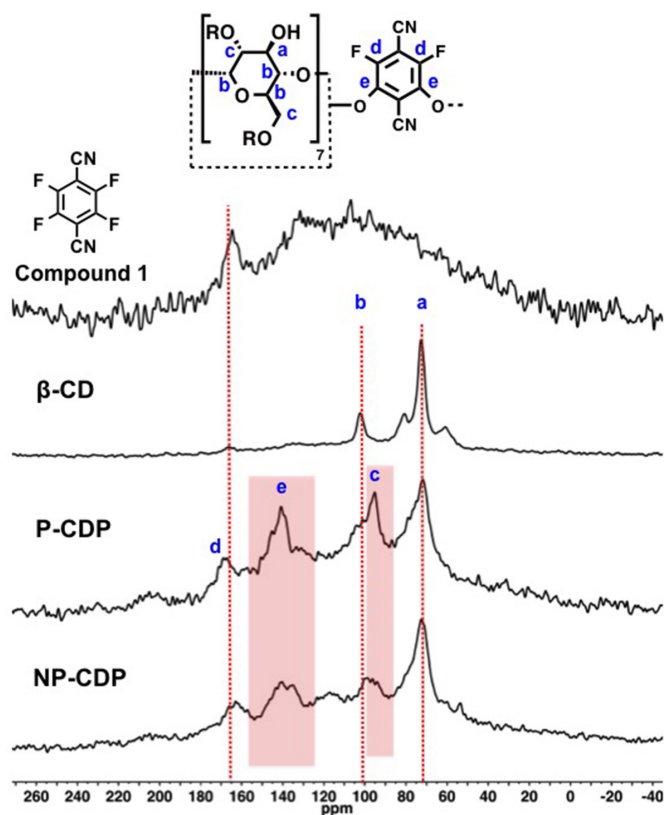


Extended Data Figure 1 | Porosity measurements of commercial ACs.

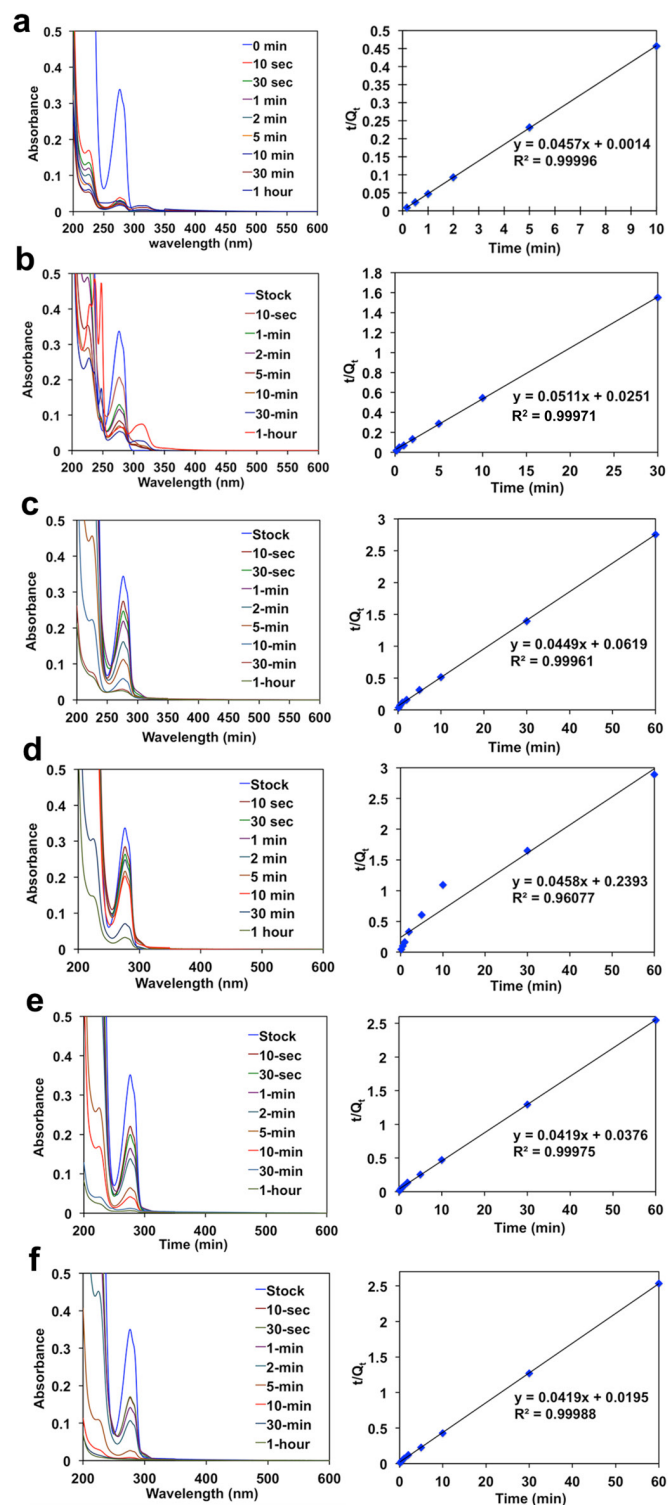
These are the materials used in Fig. 2. Shown are the N_2 sorption isotherm (77 K, left column) and the cumulative pore size distribution (right column) of Brita AC (a), GAC (b) and NAC (c). The cumulative pore size distributions of each adsorbent are similar to that of P-CDP (Fig. 1c).



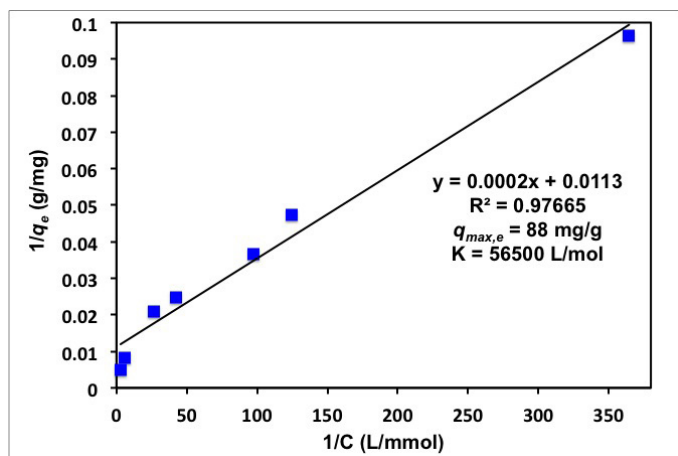
Extended Data Figure 2 | Infrared spectra of the cyclodextrin polymers and monomers. Spectra are labelled by chemical structure or compound name (top trace is 1, second trace down is β -CD). The FT-IR spectra shown in this figure of P-CDP and NP-CDP reflect the incorporation of β -CD and 1.



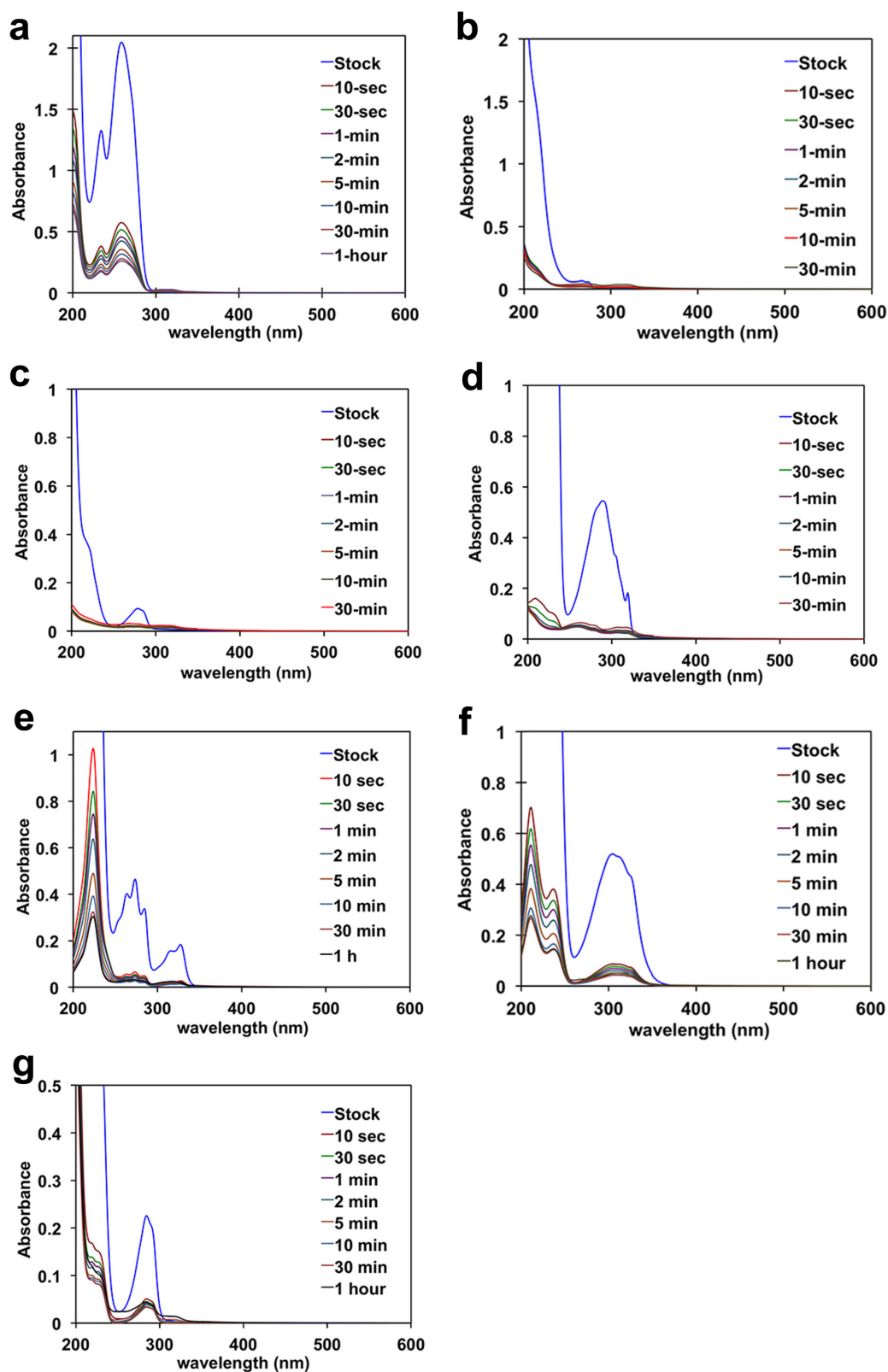
Extended Data Figure 3 | ^{13}C CP-MAS solid-state NMR spectra of P-CDP, NP-CDP, β -CD and 1. The spectra of P-CDP and NP-CDP exhibit resonances associated with β -CD at $\delta = 72$ and 100 p.p.m. (labelled a and b, respectively). Resonances at $\delta = 95$ and 140 p.p.m. (labelled e and c) correspond to the newly formed alkoxy groups and aromatic carbons, respectively. The spectrum of 1 is broadened because of ^{19}F – ^{13}C coupling.



Extended Data Figure 4 | Characterization of the bisphenol A uptake rate by each adsorbent. UV-vis spectra recorded at different contact times (coloured traces; left column) and pseudo-second-order plots (right column) for P-CDP (a), NP-CDP (b), EPI-CDP (c), Brita AC (d), GAC (e) and NAC (f). t (in min) is the contact time of bisphenol A solution with the adsorbent, and Q_t (in mg g^{-1}) is the amount of bisphenol A adsorbed per gram of adsorbent.



Extended Data Figure 5 | Langmuir isotherm of bisphenol A adsorption by P-CDP. The equilibrium uptake of bisphenol A, q_e (in mg g^{-1}), by P-CDP as a function of bisphenol A residual concentration (C , in mol l^{-1}) fits the Langmuir model, which is consistent with the formation of 1:1 inclusion complexes with an association constant (K) of $56,500 \text{ L mol}^{-1}$, and an 88 mg g^{-1} maximum equilibrium adsorption capacity ($q_{\max,e}$).



Extended Data Figure 6 | Uptake of other pollutants by P-CDP. UV-vis spectra recorded as a function of contact times with P-CDP (1 mg ml^{-1}). **a**, BPS (0.1 mM); **b**, metolachlor (0.1 mM); **c**, ethinyl oestradiol (0.04 mM); **d**, propranolol hydrochloride (0.09 mM); **e**, 2-NO (0.1 mM); **f**, 1-NA (0.1 mM); and **g**, 2,4-DCP (0.1 mM).

Extended Data Table 1 | Water regain analysis of P-CDP and NP-CDP

	Dry pore volume (cm ³ /g)	H ₂ O regain (wt%)	Volumetric H ₂ O regain (cm ³ /g)
P-CDP	0.112	265	2.65
NP-CDP	0.003	86	0.86

P-CDP takes up more water than NP-CDP, yet the latter swells to a greater degree, presumably because of its decreased crosslinking density.

Extended Data Table 2 | Rates of bisphenol A uptake by each adsorbent

Sorbent	k_{obs} (g/mg min)	Correlation coefficient R^2	Time to reach equilibrium (min)
P-CDP	1.5	1.00	10
NP-CDP	0.10	1.00	30
EPI-CDP	0.032	1.00	>60
NAC	0.090	1.00	10
GAC	0.047	1.00	>60
Brita AC	0.0088	0.96	>60

Sorbent	% Uptake in 10 sec	% Uptake at equilibrium	% Equilibrium in 10 sec ^b	q_e ^c
P-CDP	89	95	94	22
NP-CDP	39	84	46	19
EPI-CDP	20	93 ^a	22	22 ^a
NAC	52	99	53	24
GAC	37	98 ^a	38	24 ^a
Brita AC	16	90 ^a	18	21 ^a

Top: A comparison of the apparent second-order rate constants k_{obs} of the bisphenol A (initial [BPA] = 0.1 mM) uptake by each adsorbent (1 mg ml^{-1}), correlation coefficients (R^2) of the fit to the pseudo-second-order kinetic model, and the required contact time (in minutes) to reach equilibrium. These data indicate that P-CDP removes bisphenol A with an apparent second-order rate constant k_{obs} that is 15–200 times higher than the other adsorbents. P-CDP reaches 95% of its equilibrium bisphenol A uptake within 10 s. Bottom: Comparison of percentage bisphenol A removal efficiency by each adsorbent after 10 s and at equilibrium, percentage equilibrium removal efficiency after 10 s, and the amount q_e in units of milligrams per gram, of bisphenol A adsorbed by each adsorbent at equilibrium.

^aThis assumes that the equilibrium value is reached within 60 min.

^b% equilibrium value that is achieved in 10 s.

^cAmount of BPA adsorbed at equilibrium (mg BPA per g adsorbent).

Extended Data Table 3 | Equilibrium uptake of each pollutant by P-CDP

Pollutant	% Uptake at equilibrium ^a	q_e (mg g ⁻¹) ^b
BPA	95	22
BPS	85	21
metolachlor	92	26
ethinyl estradiol	89	22
propranolol	96	26
2-NO	91	14
1-NA	92	13
DCP	85	14

The uptake percentage and amount of bound micropollutant (q_e) were determined from the data shown in Extended Data Fig. 6.

^aDetermined using the % uptake value after 10 min from the adsorption kinetic studies.

^bAmount of BPA adsorbed at equilibrium (mg BPA per g adsorbent).

Extended Data Table 4 | Adsorption equilibrium constants for each micropollutant onto P-CDP

Pollutant	K (M⁻¹)^a
BPA	56,500^b
BPS	17,000^c
metolachlor	35,000^c
ethinyl estradiol	46,000^c
propranolol hydrochloride	72,000^c
2-NO	30,000^c
1-NA	35,000^c
DCP	17,000^c

The adsorption equilibrium constant *K* for bisphenol A uptake was determined using a Langmuir adsorption isotherm (Extended Data Fig. 5). The adsorption equilibrium constants for the other micropollutants were estimated from the equilibrium uptake values observed in the kinetic studies (Extended Data Fig. 6).

^aAdsorption equilibrium constant.

^bCalculated from Langmuir adsorption isotherm (Methods section 'Thermodynamic studies of adsorption').

^cEstimated from the equilibrium uptake observed at late time points of kinetic adsorption experiments.

Iron-catalysed tritiation of pharmaceuticals

Renyuan Pony Yu¹, David Hesk², Nelo Rivera², István Pelczer¹ & Paul J. Chirik¹

A thorough understanding of the pharmacokinetic and pharmacodynamic properties of a drug in animal models is a critical component of drug discovery and development^{1–6}. Such studies are performed *in vivo* and *in vitro* at various stages of the development process—ranging from preclinical absorption, distribution, metabolism and excretion (ADME) studies to late-stage human clinical trials—to elucidate a drug molecule's metabolic profile and to assess its toxicity². Radiolabelled compounds, typically those that contain ¹⁴C or ³H isotopes, are one of the most powerful and widely deployed diagnostics for these studies^{4,5}. The introduction of radiolabels using synthetic chemistry enables the direct tracing of the drug molecule without substantially altering its structure or function. The ubiquity of C–H bonds in drugs and the relative ease and low cost associated with tritium (³H) make it an ideal radioisotope with which to conduct ADME studies early in the drug development process^{2,4,6}. Here we describe an iron-catalysed method for the direct ³H labelling of pharmaceuticals by hydrogen isotope exchange, using tritium gas as the source of the radioisotope. The site selectivity of the iron catalyst is orthogonal to currently used iridium catalysts and allows isotopic labelling of complementary positions in drug molecules, providing a new diagnostic tool in drug development.

Deuterium- and tritium-labelled compounds find widespread application in the pharmaceutical industry because of their ability to alter metabolism^{7,8}, to exploit kinetic isotope effects⁹ and most commonly to perform ADME studies required for registration^{1–6}. The most common methods for preparation of ³H-radiolabelled drug molecules include reduction of a functional group to a C–³H bond using stoichiometric reagents or by catalytic methods involving direct hydrogen isotope exchange (HIE)^{2,4,6}. With the latter approach, hydrogen atoms in the drug molecule or an advanced intermediate undergo exchange with tritium gas (³H₂) or tritiated water (³H₂O) in the presence of a transition metal catalyst. Whereas both ³H₂ and ³H₂O have been applied in drug discovery and ADME studies, ³H₂ is preferred over ³H₂O for three reasons: the latter is prepared from tritium gas, is known to undergo decomposition to its constituent elements by autoradiolysis, and presents challenges to safe handling owing to a relatively high radioactivity-to-volume ratio requiring commercial sources to be diluted with natural abundance water⁶. The higher isotopic purity and atom efficiency associated with ³H₂ enables the preparation of radiolabelled drug compounds with higher specific activity and further demonstrates the preference for tritium gas for HIE. Because many known precious metal catalysts used for HIE of organic substrates and drug molecules rely on water^{6,10,11} or organic molecules^{12,13} as the source of the hydrogen isotope, the discovery of new catalyst technology that enables HIE using ³H₂ gas would be transformative for ADME studies⁶.

Among precious metal catalysts used for isotopic labelling of pharmaceuticals, Crabtree's iridium catalyst, [(COD)Ir(py)PCy₃]PF₆ (COD = 1,5-cyclooctadiene; py = pyridine), a compound initially developed for alkene hydrogenation¹⁴, and Kerr's iridium-carbene catalysts¹⁵ are the most widely used (Fig. 1)^{15–18}. These catalysts and other Ir(I) variants typically operate in a limited range of solvents, and

over the past two decades efforts have been focused on improving their stability and compatibility with functional groups commonly present in pharmaceuticals^{12,15,16,18,19}. The site selectivity of C–H exchange is predictable and relies on directing groups to enable reactivity of the *ortho* positions in aromatic and heteroaromatic rings. Depending on the specific target and purpose of the metabolic studies, isotopologues and isotopomers of the radiolabelled drug molecules—where the radiolabels are introduced at different locations in the molecule and in different amounts—may be required^{2,6}. Accordingly, new catalysts that introduce tritium at sites orthogonal to those accessed by iridium and with high degrees of incorporation and hence specific activities are valuable in augmenting existing technologies and in providing new diagnostics for ADME studies.

The bis(arylimidazol-2-ylidene)pyridine iron bis(dinitrogen) complex²⁰ has recently been shown by our laboratory to be an exceptionally active base metal catalyst for the hydrogenation of unactivated, essentially unfunctionalized alkenes^{21,22}. *In situ* monitoring of the variant bearing saturated N-heterocyclic carbenes, (H₄-iPrCNC)Fe(N₂)₂ (H₄-iPrCNC = 2,6-(2,6-ⁱPr₂-C₆H₃-4,5-H₂-imidazol-2-ylidene)₂C₅H₃N), **1**, during the course of alkene hydrogenation in perdeuterated benzene demonstrated C–H isotopic exchange between the *sp*² C–H bonds of the solvent and free H₂ gas. Quantitative ¹³C{¹H} NMR spectroscopy (126 MHz) established formation of C₆H_xT_{6–x} isotopologues with diagnostic chemical shifts (Δδ = 37–65 Hz) and multiplicity (δ_{benzene-d6} = 128.06 p.p.m., ²J_{C–2H} = 24 Hz, see Supplementary Information for representative spectrum). To explore the generality and the potential for tritium exchange reactions, additional representative substrates were evaluated with the iron catalyst.

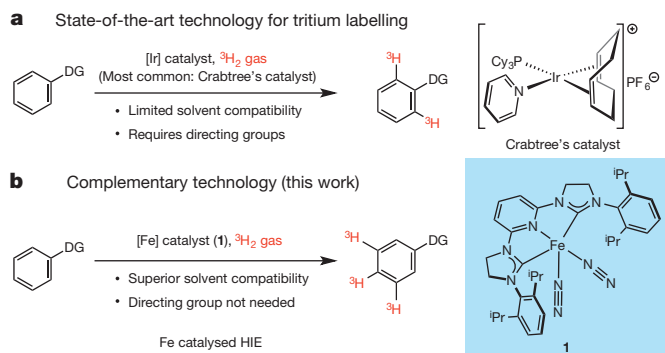


Figure 1 | Homogeneous transition metal-catalysed hydrogen/tritium exchange using tritium gas. **a**, Ir catalysts are widely used for the tritium labelling of pharmaceuticals; these catalysts operate through directed *ortho* exchange in the presence of a directing group (DG); Crabtree's catalyst (right) is the most popular and only operates in CH₂Cl₂. More recently, Kerr and co-workers developed a class of Ir-carbene variants that exhibit superior catalyst stability, efficiency and functional group compatibility^{15,18}. **b**, Complementary technology: iron catalyst (**1**) as an alternative tritiation catalyst that exhibits orthogonal C–H bond selectivity that is determined by steric accessibility and C–H bond acidity. Directing groups are not required and the catalyst operates in a range of solvents including THF, DMF and NMP.

¹Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA. ²Merck Research Laboratories, Rahway, New Jersey 07065, USA.

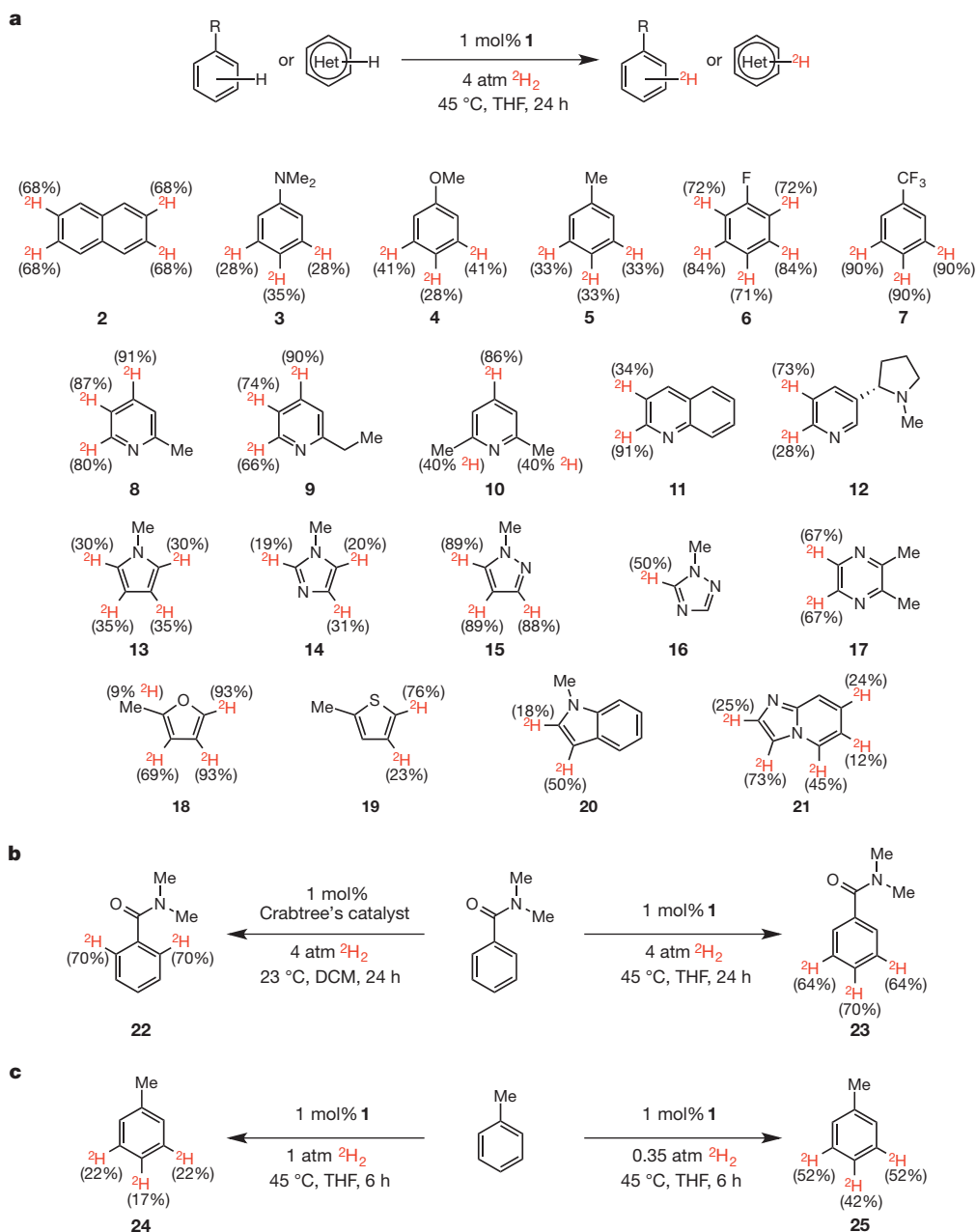


Figure 2 | Fe-catalysed deuterium labelling of representative arenes and heteroarenes. a, Here selectivity for isotopic exchange is governed by steric accessibility; percentages in parentheses correspond to the extent of deuterium incorporation at the designated positions; for **12** and **19**,

3 mol% catalyst loading was employed. **b,** Selectivity comparison between **1** and Crabtree's iridium catalyst. **c,** Inverse pressure dependence exhibited by **1**.

Catalyst evaluation studies were initially conducted with $^2\text{H}_2$ gas as a more convenient and readily handled $^3\text{H}_2$ surrogate. The iron-catalysed HIE of arenes and heteroarenes was initially examined (Fig. 2). All of the catalytic reactions were conducted under standard conditions employing 1 mol% of **1**, 3.02 M substrate in THF with 4 atm of $^2\text{H}_2$ at 45 °C. The extent of isotopic exchange was determined after 24 h using a combination of quantitative ^1H and quantitative $^{13}\text{C}\{^1\text{H}\}$ NMR spectroscopies. The latter analytical technique has been especially informative for substrates where multiple aromatic C–H signals are poorly resolved in ^1H NMR spectra. Iron pre-catalyst **1** has proven particularly effective for the $^1\text{H}/^2\text{H}$ exchange with many important substructures in pharmaceuticals, including substituted arenes (**2–7**), pyridine derivatives (**8–12**) as well as other nitrogen-, oxygen- and sulphur-containing heteroarenes (**13–21**)^{23,24}. Examination of the site selectivity of the iron-catalysed deuteration of arenes (**2–7**) established

two salient trends—namely, electron-poor substrates (**6**, **7**) undergo C–H isotopic exchange faster than electron-rich ones (for example, **3–5**), and the site selectivity of the C–H bond activation occurs at the most sterically accessible C–H bonds. The iron-catalysed method is also versatile within various classes of nitrogen heterocycles. With (–)-nicotine (**12**) exclusive deuteration of the sterically accessible positions of the pyridine occur in the presence of the *N*-methyl pyrrolidine. With 2,6-lutidine (**10**), the 4-position is the only sp^2 hybridized C–H bond to undergo isotopic exchange. Notably, significant (40%) deuterium incorporation was observed in the sp^3 benzylic positions, a feature absent in 2-methylpyridine (**8**) and 2-ethylpyridine (**9**). In the case of *N*-methyl-indole (**20**) and imidazo[1,2- α]pyridine (**21**), exchange occurred at positions adjacent to the ring junctions.

The complementary selectivity of **1** as compared to Crabtree's catalyst is highlighted by the deuteration of *N,N*-dimethylbenzamide (**22**).

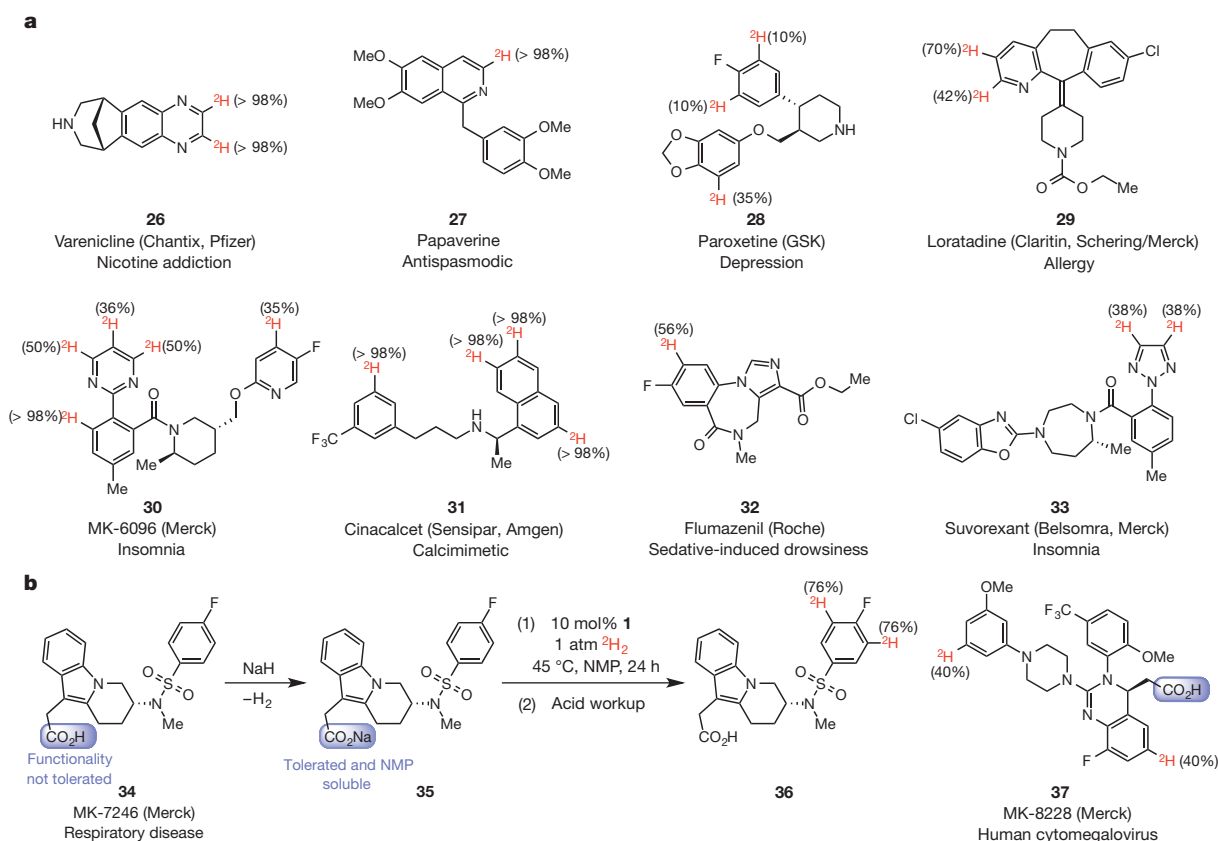


Figure 3 | Fe-catalysed deuterium labelling of drug molecules.

a, Direct labelling of drug substrates, general catalytic reaction conditions as follows: 10 mol% **1**, 0.154 mmol substrate, 0.5 ml NMP solution, 1 atm $^2\text{H}_2$, 45 °C, 24 h. **b**, Deprotonation-protection strategy of carboxylic acids

As shown in Fig. 2b, the iridium catalyst promoted isotopic exchange exclusively at the *ortho* positions as expected for an amide-directed C–H activation. In contrast, the iron catalyst enabled deuterium incorporation statistically at the sterically accessible *meta* and *para* positions. Precious metal catalysts exhibiting sterically preferred C–H selectivity are well documented, although in most cases $^2\text{H}_2\text{O}$ or organic solvents serve as the source of the isotopic label^{10,11,25}. An iridium example, $[(\eta^5\text{-C}_5\text{Me}_5)(\text{Me}_3\text{P})\text{Ir}(\text{CH}_3)_2\text{CH}_2\text{Cl}_2][\text{BAr}^{\text{F}}_4]$ (where $[\text{BAr}^{\text{F}}_4] = \text{tetrakis}(3,5\text{-bis(trifluoromethyl)phenyl)borate}$) has been reported whose selectivity is determined by sterically accessible C–H bonds and employs $^2\text{H}_2$ or $^3\text{H}_2$ gas but requires a stoichiometric rather than catalytic amount of metal complex²⁵. Because practical tritiation of pharmaceuticals are typically conducted with a modest excess of tritium gas (that is, at subatmospheric pressure)⁶, the deuteration of toluene catalysed by **1** was evaluated at both 1 atm and 0.35 atm of $^2\text{H}_2$ gas. Higher activity was observed at the lower pressures of $^2\text{H}_2$ gas, as 42% and 52% deuterium incorporation was observed in the *para* and *meta* positions after 6 h. These values were reduced to 17% and 22%, respectively, when the pressure was increased to 1 atm (Fig. 2c).

The promising functional group compatibility and the favourable inverse pressure dependence observed with $^1\text{H}/^2\text{H}$ exchange with **1** prompted evaluation of the deuteration and tritiation of pharmaceuticals (Fig. 3). Although THF was used for catalyst evaluation studies with representative arenes and heteroarenes, the limited solubility of more functionalized pharmaceuticals in this solvent inspired use of other media for catalytic HIE. Iron catalyst **1** has proven robust and operated efficiently in a range of polar aprotic solvents including dimethylacetamide (DMA), dimethylformamide (DMF) and N-methylpyrrolidinone (NMP)—the last being used for all subsequent drug labelling studies. With 10 mol% **1**, selective deuteration of papaverine (**27**), an opium alkaloid antispasmodic²⁶, was achieved where

present in **34** and **37** for compatibility with the iron catalyst. For **37**, the same deprotonation strategy was employed, but the carboxylate salt was generated *in situ* and used without further isolation.

exclusive deuterium incorporation was observed at the 3-position of the isoquinoline fragment. This site selectivity was as expected on the basis of the small molecule studies—this C–H bond is the only sterically accessible position for the iron catalyst. Other commercially available pharmaceuticals, namely varenicline, loratadine, cinacalcet and flumazenil, underwent iron-catalysed deuteration at the predicted sterically accessible sites. With loratadine, isotopic exchange was preferred at the 3-position over the 2-position of the ring, contrary to typical C–H functionalization preferences. With the experimental orexin receptor antagonist, MK-6096 (**30**)²⁷, quantitative deuterium incorporation occurred at the 3-position of the 1,2,5-trisubstituted benzene subunit, probably through *ortho*-directed C–H activation from the nitrogen atom of the proximal pyrimidine ring. This preferred selectivity demonstrates that in the presence of appropriate ligands, directed C–H activation can be preferential to steric site selectivity. Pharmaceuticals such as the CRTH2 antagonist, MK-7246 (**34**)²⁸, and the anti-viral drug candidate, MK-8228²⁹, proved incompatible with **1**, probably as a result of the carboxylic acid functional group. Preparation of the NMP soluble conjugate sodium carboxylate salt of MK-7246, **35**, readily obtained by addition of NaH to the free acid, overcame this limitation. Following acidic workup after iron-catalysed deuteration provided the isotopically enriched product, **36**, in high yield with deuterium selectively incorporated into the C–H bonds *ortho* to the fluorine. As illustrated with MK-8228 (**37**), a Merck developmental drug candidate for treatment of Human Cytomegalovirus (HCMV) infections in transplant patients currently undergoing Phase III clinical trials, the carboxylate salt was generated *in situ* and used in the HIE reaction without isolation or further purification.

To demonstrate the utility and to further highlight the complementarity of the iron catalyst with existing iridium-catalysed methods, a series of representative, commercially available pharmaceuticals were

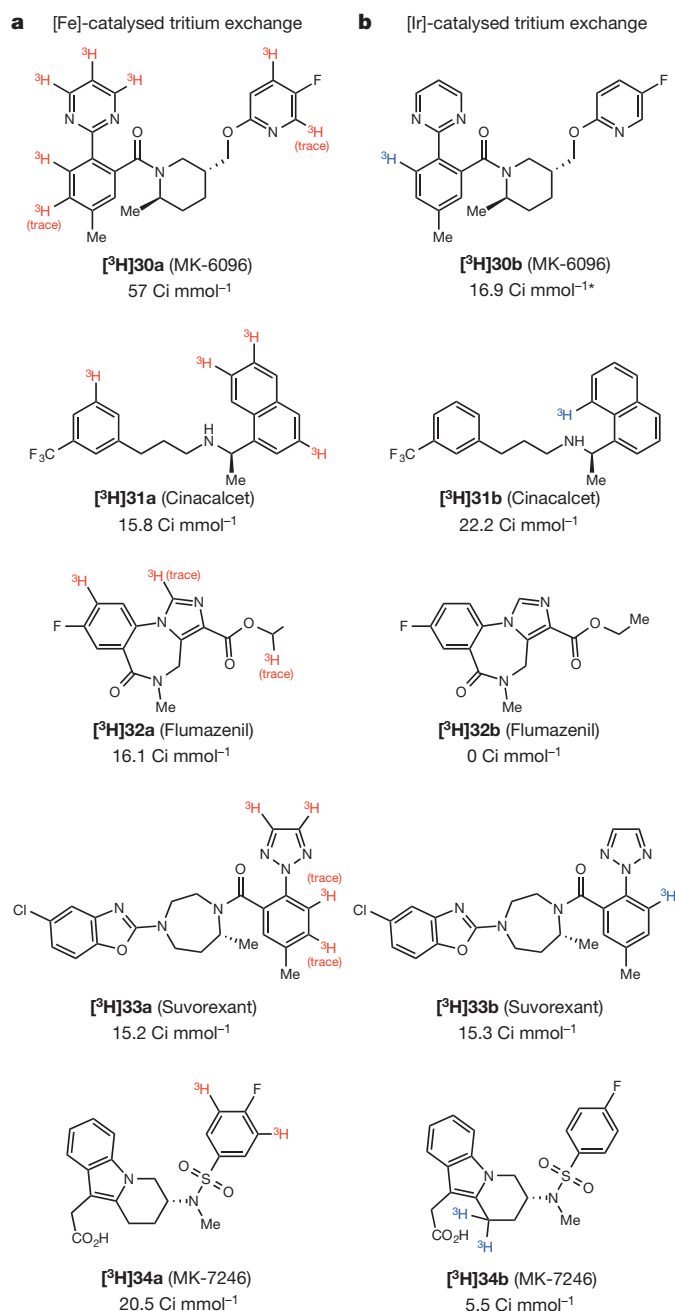


Figure 4 | Tritium labelling of drug molecules. **a**, Using **1**, reaction conditions as follows: 25 mol% catalyst loading, 7 μ mol substrate, 1.2 Ci $^3\text{H}_2$ (0.15 atm), 0.2 ml NMP, 23 °C, 16 h. **b**, Using Crabtree's catalyst, reaction conditions as follows: 25 mol% catalyst loading, 7 μ mol substrate, 1.2 Ci tritium (0.15 atm), 0.5 ml CH_2Cl_2 , 23 °C, 16 h. For [³H]**34a**, the sodium salt conjugate base, **35** was used in place of **34** owing to incompatibility of the carboxylic acid functionality with **1** (see Supplementary Information for details). Specific activities for each compound are given in Ci mmol⁻¹. *A comparison using another Ir-based catalyst, [(COD)Ir(IMes)PPh₃]₂PF₆ (see Supplementary Information for catalyst representation), originally developed by Kerr and co-workers^{15,18}, was performed using MK-6096 as the model substrate; the Kerr catalyst labelled the same position as Crabtree's catalyst with a lower specific activity of 8.8 Ci mmol⁻¹ under identical conditions.

tritiated using both **1** (Fig. 4a) and Crabtree's iridium catalyst (Fig. 4b). Use of $^3\text{H}_2$ gas was particularly informative, to not only determine site selectivity but to also establish the relative specific activity of the final product between the base and precious metal C–H exchange catalysts. In all cases, tritium exchange reactions were performed using

subatmospheric pressures (~ 0.15 atm, approximately 12 equiv. ^3H atoms with respect to substrate) of $^3\text{H}_2$ gas at room temperature at higher (25 mol%) catalyst loading, and the positions of labelling were confirmed by ^3H NMR spectroscopy. With MK-6096 ([³H]**30a**), flumazenil ([³H]**32a**) and suvorexant ([³H]**33a**), tritium was detected in trace quantities from iron-catalysed isotopic exchange in positions previously unobserved in the deuteration experiments (Fig. 4a). These positions are adjacent to arene substitution such as methyl groups or ring junctions where C–H activation has a higher barrier due to steric inaccessibility. The high sensitivity of ^3H NMR spectroscopy as compared to the corresponding ^1H and ^{13}C experiments allowed detection of the small quantities of isotopic label introduced into these positions. For suvorexant ([³H]**33a**), the iron and iridium exchange methods are complementary; the iron catalyst prefers isotopic exchange at the relatively electron deficient triazole, whereas iridium favours directed C–H functionalization and tritium incorporation in the aryl ring at the site *ortho* to the triazole subunit. Similar orthogonal yet complementary site selectivity was observed with MK-7246 where after carboxylate deprotonation, the iron catalyst enables sterically driven tritiation of the fluorinated arene ring while iridium promoted directed C–H exchange exclusively in the saturated nitrogen heterocycle.

The tritiation of flumazenil highlights the beneficial reactivity enabled by **1**. Sterically accessible C–H bonds *ortho* to fluorine in the arene fragment resulted in a relatively high specific activity of 16.1 Ci mmol⁻¹ ([³H]**32a**). Typically, specific activity values in the range 10–20 Ci mmol⁻¹ are acceptable for ADME studies³⁰. With Crabtree's iridium catalyst, no tritiation of flumazenil was observed under similar conditions (Fig. 4b, [³H]**32b**). Although amide and ester functionalities are present that could potentially serve as directing groups and hence enable isotopic exchange, conformational effects probably inhibit the accessibility of proximal C–H bonds required for tritiation. Significantly higher specific activity was observed with the iron-catalysed tritiation of MK-6096. This difference may be traced to the abundance of sterically accessible C–H bonds where the iridium catalyst only introduces the isotopic label at one site (Fig. 4b, [³H]**30b**).

In summary, an iron catalyst for the tritiation of pharmaceuticals has been discovered with selectivity for C–H bonds that is orthogonal and complementary to existing precious metal catalyst technology. This catalyst is compatible with a range of pharmaceutically relevant functional groups and operates efficiently in polar aprotic solvents at low pressures of tritium gas. The ability to introduce radiolabels into unique positions provides a new diagnostic for ADME studies, a critical component of the drug approval process. Current efforts are devoted to elucidating the mechanism of action and to improve the stability and handling of the iron precursor.

Received 29 August; accepted 17 November 2015.

- Lappin, G. & Temple, S. *Radiotracers in Drug Development* (CRC Press, 2006).
- Isin, E. M., Elmore, C. S., Nilsson, G. N., Thompson, R. A. & Weidolf, L. Use of radiolabeled compounds in drug metabolism and pharmacokinetic studies. *Chem. Res. Toxicol.* **25**, 532–542 (2012).
- Marathe, P. H., Shyu, W. C. & Humphreys, W. G. The use of radiolabeled compounds for ADME studies in discovery and exploratory development. *Curr. Pharm. Des.* **10**, 2991–3008 (2004).
- Lockley, W. J. S., McEwen, A. & Cooke, R. Tritium: a coming of age for drug discovery and development ADME studies. *J. Labelled Comp. Radiopharm.* **55**, 235–257 (2012).
- Elmore, C. S. The use of isotopically labeled compounds in drug discovery. *Annu. Rep. Med. Chem.* **44**, 515–534 (2009).
- Voges, R., Heys, J. R. & Moenius, T. *Preparation of Compounds Labeled with Tritium and Carbon-14* (John Wiley, 2009).
- Meanwell, N. A. Synopsis of some recent tactical application of bioisosteres in drug design. *J. Med. Chem.* **54**, 2529–2591 (2011).
- Katsnelson, A. Heavy drugs draw heavy interest from pharma backers. *Nature Med.* **19**, 656 (2013).
- Jarman, M. *et al.* The deuterium isotope effect for the α -hydroxylation of tamoxifen by rat liver microsomes accounts for the reduced genotoxicity of [D₅-ethyl]tamoxifen. *Carcinogenesis* **16**, 683–688 (1995).
- Atzrodt, J., Derdau, V., Fey, T. & Zimmermann, J. The renaissance of H/D exchange. *Angew. Chem. Int. Edn* **46**, 7744–7765 (2007).

11. Klei, S. R., Golden, J. T., Tilley, T. D. & Bergman, R. G. Iridium-catalyzed H/D exchange into organic compounds in water. *J. Am. Chem. Soc.* **124**, 2092–2093 (2002).
12. Ma, S., Villa, G., Thuy-Boun, P. S., Homs, A. & Yu, J.-Q. Palladium-catalyzed ortho-selective C-H deuteration of arenes: evidence for superior reactivity of weakly coordinated palladacycles. *Angew. Chem. Int. Edn* **53**, 734–737 (2014).
13. Zhou, J. & Hartwig, J. F. Iridium-catalyzed H/D exchange at vinyl groups without olefin isomerization. *Angew. Chem. Int. Edn* **47**, 5783–5787 (2008).
14. Crabtree, R., Felkin, H. & Morris, G. Cationic iridium diolefin complexes as alkene hydrogenation catalysts and isolation of some related hydrido complexes. *J. Organomet. Chem.* **141**, 205–215 (1977).
15. Nilsson, G. N. & Kerr, W. J. The development and use of novel iridium complexes as catalysts for ortho-directed hydrogen isotope exchange reactions. *J. Labelled Comp. Radiopharm.* **53**, 662–667 (2010).
16. Hesk, D., Das, P. R. & Evans, B. Deuteration of acetanilides and other substituted aromatics using $[\text{Ir}(\text{COD})(\text{Cy}_3\text{P})(\text{Py})]\text{PF}_6$ as catalyst. *J. Labelled Comp. Radiopharm.* **36**, 497–502 (1995).
17. Lockley, W. J. S. & Heys, J. R. Metal-catalysed hydrogen isotope exchange labelling: a brief overview. *J. Labelled Comp. Radiopharm.* **53**, 635–644 (2010).
18. Brown, J. A. *et al.* The synthesis of highly active iridium(I) complexes and their application in catalytic hydrogen isotope exchange. *Adv. Synth. Catal.* **356**, 3551–3562 (2014).
19. Salter, R. The development and use of iridium(I) phosphine systems for ortho-directed hydrogen-isotope exchange. *J. Labelled Comp. Radiopharm.* **53**, 645–657 (2010).
20. Danopoulos, A. A., Wright, J. A. & Motherwell, W. B. Molecular N_2 complexes of iron stabilised by *N*-heterocyclic ‘pincer’ dicarbene ligands. *Chem. Commun.* 784–786 (2005).
21. Yu, R. P. *et al.* High-activity iron catalysts for the hydrogenation of hindered, unfunctionalized alkenes. *ACS Catal.* **2**, 1760–1764 (2012).
22. Yu, R. P. *et al.* Catalytic hydrogenation activity and electronic structure determination of bis(arylimidazol-2-ylidene)pyridine cobalt alkyl and hydride complexes. *J. Am. Chem. Soc.* **135**, 13168–13184 (2013).
23. Vitaku, E., Smith, D. T. & Njardarson, J. T. Analysis of the structural diversity, substitution patterns, and frequency of nitrogen heterocycles among U.S. FDA approved pharmaceuticals. *J. Med. Chem.* **57**, 10257–10274 (2014).
24. Taylor, R. D., MacCoss, M. & Lawson, A. D. G. Rings in drugs. *J. Med. Chem.* **57**, 5845–5859 (2014).
25. Skaddan, M. B., Yung, C. M. & Bergman, R. G. Stoichiometric and catalytic deuterium and tritium labeling of ‘unactivated’ organic substrates with cationic $\text{Ir}(\text{III})$ complexes. *Org. Lett.* **6**, 11–13 (2004).
26. Liu, J. K. & Couldwell, W. T. Intra-arterial papaverine infusions for the treatment of cerebral vasospasm induced by aneurysmal subarachnoid hemorrhage. *Neurocrit. Care* **2**, 124–132 (2005).
27. Coleman, P. J. *et al.* Discovery of $[(2R,5R)\text{-}5\text{-}[[[5\text{-fluoropyridin-2-yl}]\text{oxy}]\text{methyl}]\text{-}2\text{-methylpiperidin-1-yl}][5\text{-methyl-}2\text{-}(\text{pyrimidin-2-yl})\text{phenyl}]\text{methanone}$ (MK-6096): a dual orexin receptor antagonist with potent sleep-promoting properties. *ChemMedChem* **7**, 415–424 (2012).
28. Molinaro, C. *et al.* CRTH2 antagonist MK-7246: a synthetic evolution from discovery through development. *J. Org. Chem.* **77**, 2299–2309 (2012).
29. Lischka, P. *et al.* *In vitro* and *in vivo* activities of the novel anticytomegalovirus compound AIC246. *Antimicrob. Agents Chemother.* **54**, 1290–1297 (2010).
30. Nagasaki, T. *et al.* A new practical tritium labelling procedure using sodium borotritide and tetrakis(triphenylphosphine)palladium(0). *J. Labelled Comp. Radiopharm.* **44**, 993–1004 (2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements Merck and the Intellectual Property Accelerator Fund at Princeton University are acknowledged for financial support. We thank M. Tudge, I. Mergelsberg, L.-C. Campeau and I. Davies for discussions. We also thank D. Schenk and Y. Liu for assistance in ^3H NMR assignments.

Author Contributions R.P.Y. and P.J.C. discovered the iron-catalysed reaction. R.P.Y. performed initial deuterium exchange studies. R.P.Y. and I.P. performed the analysis of deuterium labelled products. I.P. developed and implemented the quantitative ^{13}C NMR protocol for analysis of deuterium labelled products. R.P.Y., D.H. and N.R. performed and analysed tritium-labelling studies. R.P.Y., D.H. and P.J.C. prepared the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.J.C. (pchirik@princeton.edu).

Critical insolation–CO₂ relation for diagnosing past and future glacial inception

A. Ganopolski¹, R. Winkelmann^{1,2} & H. J. Schellnhuber^{1,3}

The past rapid growth of Northern Hemisphere continental ice sheets, which terminated warm and stable climate periods, is generally attributed to reduced summer insolation in boreal latitudes^{1–3}. Yet such summer insolation is near to its minimum at present⁴, and there are no signs of a new ice age⁵. This challenges our understanding of the mechanisms driving glacial cycles and our ability to predict the next glacial inception⁶. Here we propose a critical functional relationship between boreal summer insolation and global carbon dioxide (CO₂) concentration, which explains the beginning of the past eight glacial cycles and might anticipate future periods of glacial inception. Using an ensemble of simulations generated by an Earth system model of intermediate complexity constrained by palaeoclimatic data, we suggest that glacial inception was narrowly missed before the beginning of the Industrial Revolution. The missed inception can be accounted for by the combined effect of relatively high late-Holocene CO₂ concentrations and the low orbital eccentricity of the Earth⁷. Additionally, our analysis suggests that even in the absence of human perturbations no substantial build-up of ice sheets would occur within the next several thousand years and that the current interglacial would probably last for another 50,000 years. However, moderate anthropogenic cumulative CO₂ emissions of 1,000 to 1,500 gigatonnes of carbon will postpone the next glacial inception by at least 100,000 years^{8,9}. Our simulations demonstrate that under natural conditions alone the Earth system would be expected to remain in the present delicately balanced interglacial climate state, steering clear of both large-scale glaciation of the Northern Hemisphere and its complete deglaciation, for an unusually long time.

In accordance with classical Milankovitch theory¹, interglacials—warm intervals with the lowest global ice volume—occur during periods of high summer insolation in the boreal latitudes of the Northern Hemisphere. In the past, a decrease in Northern Hemisphere insolation to below its present-day level always led to the end of interglacials and rapid growth of continental ice sheets^{2,3}, accompanied by a reduction in CO₂ concentration^{10,11}. However, at present, although summer insolation at 65° N is close to its minimum⁴, there is no evidence for the beginning of a new ice age. On the contrary, sea level, which reflects changes in global ice volume, remained essentially constant over the past several millennia^{5,12}.

The most straightforward explanation for the lack of glacial inception at present is that the current insolation minimum is not deep enough because of the low orbital eccentricity of the Earth. However, glacial inceptions have occurred in the past under similar orbital configurations. Marine Isotope Stage (MIS) 11 (about 400,000 years before present, 400 kyr BP) is often considered a close palaeo-analogue for the current interglacial (the Holocene, or MIS1) owing to the similarly low values of the eccentricity of Earth's orbit and similar CO₂ level at that time⁷ (Fig. 1). The only difference between the insolation minimum at about 400 kyr BP and the present one is a lower obliquity during MIS11.

With respect to the orbital parameters, MIS19 (about 800 kyr BP) is an even closer analogue for the Holocene (see Fig. 1). Following this analogy, it has been suggested that the current interglacial would end naturally within the next 1,500 years if the CO₂ concentration had stayed at a level of about 240 parts per million (p.p.m.), as was the case at the end of MIS19 (ref. 13). However, during the late Holocene before the beginning of the industrial era, the CO₂ concentration was about

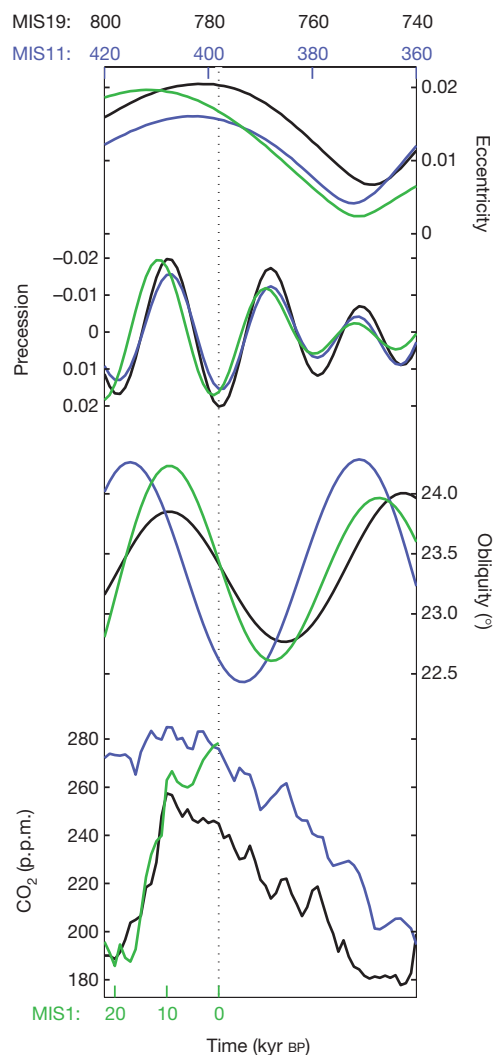


Figure 1 | Orbital parameters. Comparison of Earth's orbital parameters and CO₂ concentrations for MIS1 (green), MIS11 (blue) and MIS19 (black). The vertical dashed line corresponds to the present day for MIS1 and the minima of the precessional component of insolation for MIS11 and MIS19.

¹Potsdam Institute for Climate Impact Research, 14412 Potsdam, Germany. ²Physics Institute, Potsdam University, 14476 Potsdam, Germany. ³Santa Fe Institute, Santa Fe, New Mexico 87501, USA.

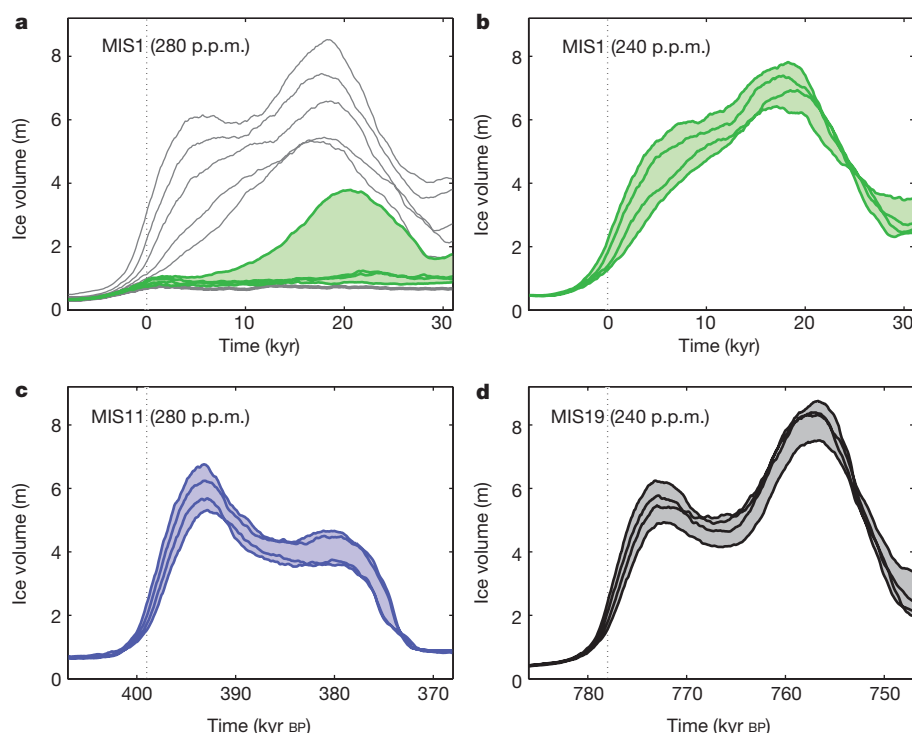


Figure 2 | Evolution of the Northern Hemisphere ice volume. Ice volume (excluding the Greenland Ice Sheet) is given in metres of sea-level equivalent for the Holocene and near future with a CO₂ concentration of 280 p.p.m. (a), the Holocene and near future with a CO₂ concentration of 240 p.p.m. (b), MIS11 with a CO₂ concentration of 280 p.p.m. (c), and MIS19 with a CO₂ concentration of 240 p.p.m. (d). Individual model

280 p.p.m., a level that is also typical for several previous interglacials. Therefore, MIS19, with its low CO₂ level, may not be a proper analogue for the present interglacial either.

A key question we address here is whether subtle differences in orbital configurations and CO₂ concentration between MIS11 and MIS19 and the Holocene are sufficient to explain the fundamentally different evolution of the climate system in the vicinity of its present state. This could have important implications for the future evolution of the Earth system.

The results presented here are based on simulations with the Earth system model of intermediate complexity CLIMBER-2 (ref. 14), which includes the three-dimensional thermomechanical ice sheet model SICOPOLIS¹⁵. The CLIMBER-2 model has been successfully applied for simulating the last eight glacial cycles¹⁶. It is known that glacial inception is associated with highly nonlinear dynamics of the climate–cryosphere system¹⁷. Previous studies^{4,18} demonstrated high sensitivity of the timing of the next glacial inception to modelling parameters and prescribed CO₂ concentration. This is consistent with the finding that glacial inception represents a bifurcation transition between interglacial and glacial climate states and that under the current orbital configuration and CO₂ concentration, the Earth system is very close to this bifurcation point^{19,20}. To ensure that the model correctly simulates the position of this bifurcation transition in the phase space of external forcings, we use the well established fact that glacial inception occurred at the end of MIS19 and MIS11, but not in the recent past. We quantify these constraints as following: for MIS19 and MIS11 only model versions that simulate a build-up of ice sheets with a total volume of at least 5 m sea-level equivalent are accepted. For the Holocene, on the contrary, all model versions simulating more than 1 m sea-level equivalent of ice growth before today are rejected. Note that these constraints are only applied to the Northern Hemisphere and do not include the Greenland Ice Sheet.

simulations from the subset of model versions compatible with the empirical constraints are shown; the shading illustrates the entire range. For MIS1 (the Holocene), with a CO₂ concentration of 280 p.p.m. (a), the model realizations that are not compatible with the observational constraints are shown as grey lines. The vertical dashed lines correspond to the minima of the precessional component of insolation.

By perturbing one of the model parameters, which affects the surface temperature in a uniform manner, we created an ensemble of twenty model realizations that differ only slightly in their unperturbed climate state (see Methods). We then applied these model realizations to simulate MIS19, MIS11 and the Holocene forced by changing orbital parameters and the prescribed CO₂ concentrations typical for each interglacial (240 p.p.m. for MIS19, and 280 p.p.m. for MIS11 and the Holocene). We selected those realizations that are compatible with the observational constraints described above. Only four model versions pass this test successfully.

None of these model versions simulate substantial ice sheet growth within the next several thousands of years under the pre-industrial CO₂ concentration of 280 p.p.m. (Fig. 2). Three of the four model versions predict the next glacial inception about 50,000 years from now, which is consistent with previous modelling results. Only the coldest model version simulates a slow growth of medium-sized ice caps in the future, mostly over the Canadian Archipelago and Scandinavia. This slow growth is attributed to decreasing obliquity, which becomes the dominant astronomical factor under the currently very low eccentricity. However, even in this case, the ice volume only crosses the glaciation threshold 20,000 years from now.

The situation is completely different for a CO₂ concentration of 240 p.p.m., which is close to that observed at the end of MIS19. In this case all four model versions simulate rapid ice growth several thousands of years before the present and large ice sheets exist already at the present time (Extended Data Fig. 1). This means that the Earth system would already be well on the way towards a new glacial state if the pre-industrial CO₂ level had been merely 40 p.p.m. lower than it was during the late Holocene, which is consistent with previous results^{13,21}.

Whether this narrow escape from glacial inception was natural remains debatable. It has been proposed²¹ that pre-industrial land-use at least partly contributed to the high Holocene CO₂ level, but the magnitude of this contribution is very uncertain^{22,23}. This escape from

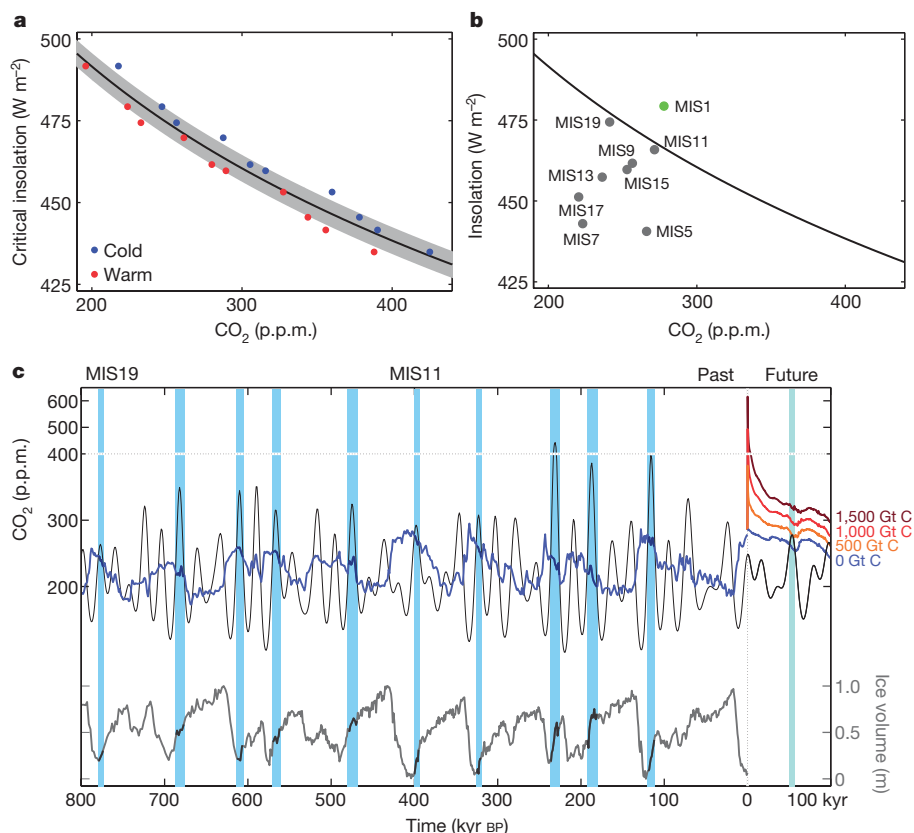


Figure 3 | Critical insolation–CO₂ relation. **a**, Best-fit logarithmic relation (black line) between the maximum summer insolation at 65° N and the CO₂ threshold for glacial inception; grey shaded area indicates ±1 s.d. Blue dots correspond to the coldest model version and red dots to the warmest. **b**, The locations of previous glacial inceptions in insolation–CO₂ phase space relative to the best-fit logarithmic curve from **a**. Glacial inception is only possible when the point is located below the insolation–CO₂ curve. **c**, The timing of past and future glacial inceptions can be explained by the CO₂ concentration and the insolation–CO₂ relation. The thin grey line depicts the CO₂ threshold value for

glacial inception, derived as a function of the maximum summer insolation at 65° N. The CO₂ concentration from ice core data^{10,11} for the past 800,000 years is shown (blue line), along with the CO₂ scenarios of 0 Gt C cumulative anthropogenic emissions (blue line), 500 Gt C (orange line), 1,000 Gt C (red line) and 1,500 Gt C (dark red line). Pale blue vertical bars indicate the time periods when the reconstructed value is below the critical CO₂ concentration, and the light blue bar shows the timing of a possible next glacial inception. The horizontal dotted line indicates the present-day CO₂ level. The lower curve depicts a proxy for the global ice volume²⁵ (thick grey line).

glacial inception in the recent past makes it very unlikely that glaciation will start in the near future without a considerable drop in CO₂ concentration. That is, glaciation did not begin even under the most favourable conditions in the past millennium for it to do so, and we do not anticipate such conditions occurring again in the foreseeable future. This is apparent from a stability analysis of the Earth system in the phase space of orbital forcing and CO₂ concentration.

Using a range of model realizations consistent with palaeoclimatic constraints, we mapped the threshold value of CO₂ leading to glacial inception depending on the maximum summer insolation at 65° N. The maximum summer insolation at 65° N is the most common metric for the orbital forcing. It has been shown that for given CO₂ concentration, the insolation threshold for glacial inception depends also on obliquity, but this dependence is rather weak¹⁹. For the sake of simplicity and consistency with our previous results, we therefore use maximum summer insolation at 65° N as the proxy for orbital forcing.

Figure 3a shows that the individual points in the insolation–CO₂ space representing different combinations of orbital parameters are clustered around the logarithmic curve. This is consistent with the fact that radiative forcing of CO₂ is proportional to the logarithm of CO₂ concentration and that in the CLIMBER-2 model, similar to many other climate models, the temperature response to CO₂ and orbital forcing is linear within the considered range of CO₂ concentrations. The critical summer insolation at 65° N can be described as $S = \alpha \ln([\text{CO}_2]/280) + \beta$, where $\alpha = -77 \text{ W m}^{-2}$ and $\beta = 466 \text{ W m}^{-2}$ and $[\text{CO}_2]$ is the concentration of CO₂ in parts per million (see Methods).

The validity of the critical insolation–CO₂ concentration relationship is confirmed for all previous glacial inceptions, because each of the past glaciations occurred when the CO₂ concentration was lower than the threshold value (Fig. 3b, c). The conditions for glacial inception are not met in the near future. For the orbital forcing of the late Holocene, the CO₂ concentration remains above the threshold value owing to the currently low eccentricity. Only in about 50,000 years from now, the threshold CO₂ value approaches the pre-industrial CO₂ concentration. Therefore, even without anthropogenic perturbations, the Holocene would be an unusually long interglacial. This also implies that the Holocene has no proper palaeoclimate analogue within at least the past million years.

Owing to the extremely long lifetime of anthropogenic CO₂ in the atmosphere, past and future anthropogenic CO₂ emissions will have a strong impact on the timing of the next glacial inception^{8,9}. In order to estimate the earliest possible onset of a new glaciation, we forced the four valid model realizations by orbital variations and different CO₂ concentration scenarios computed for the next 100,000 years with the CLIMBER-2 model (see Methods). Initial conditions were based on results from simulations of the last glacial cycle²⁴. Without anthropogenic greenhouse gas emissions, the CO₂ concentration gradually declines in the future and oscillates slightly on the orbital timescale. Note that the long-term natural future evolution of CO₂ concentration is sensitive to model parameters and is not well constrained by empirical data. Assuming that anthropogenic CO₂ emissions will be reduced to zero on a centennial timescale (not considering the possibility of negative CO₂ emissions), the long-term evolution of the CO₂

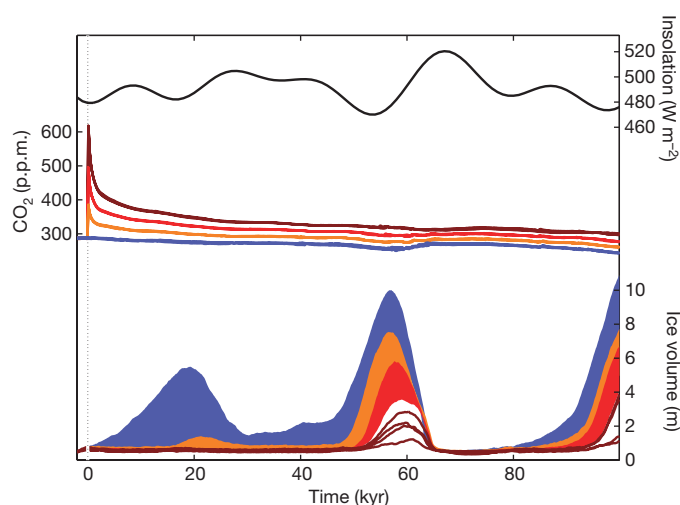


Figure 4 | The next glacial inception. The top panel shows the temporal evolution of the maximum summer insolation at 65° N. The middle panel shows the simulated CO₂ concentration during the next 100,000 years for different cumulative CO₂ emission scenarios: 0 Gt C anthropogenic emissions (blue), 500 Gt C (orange), 1,000 Gt C (red) and 1,500 Gt C (dark red line). The bottom panel shows simulated ice volume corresponding to the different CO₂ emission scenarios. Individual simulations are shown for the 1,500 Gt C scenario; for the other scenarios, the range is given as shading.

concentration depends solely on the total cumulative carbon emissions within the next centuries.

Under three scenarios with cumulative emissions of 500 gigatonnes of carbon (Gt C), 1,000 Gt C and 1,500 Gt C, we simulate the ice volume on the Northern Hemisphere for the next 100,000 years. Even for a total of 500 Gt C cumulative emissions, which is only slightly above the present-day value, the evolution of the Northern Hemisphere ice sheets is affected over tens of thousands of years (Fig. 4). In the 1,000 Gt C scenario, the probability of glacial inception during the next 100,000 years is notably reduced, and under cumulative emissions of 1,500 Gt C, glacial inception is very unlikely within the entire 100,000 years. This confirms our conclusions from the critical insolation threshold for glacial inception. Because all 2013 Intergovernmental Panel on Climate Change scenarios—except Representative Concentration Pathway 2.6 (RCP2.6), which leads to the total radiative forcing of greenhouse gases of 2.6 W m⁻² in 2100—imply that cumulative carbon emission will exceed 1,000 Gt in the twenty-first century, our results suggest that anthropogenic interference will make the initiation of the next ice age impossible over a time period comparable to the duration of previous glacial cycles.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 October 2014; accepted 19 November 2015.

1. Milanković, M. M. *Canon of Insolation and the Ice-Age Problem* (Königlich Serbische Akademie, 1941).
2. Hays, J. D., Imbrie, J. & Shackleton, N. J. Variations in the Earth's orbit: pacemaker of the Ice Ages. *Science* **194**, 1121–1132 (1976).
3. Paillard, D. The timing of Pleistocene glaciations from a simple multiple-state climate model. *Nature* **391**, 378–381 (1998).
4. Berger, A. & Loutre, M. F. An exceptionally long interglacial ahead? *Science* **297**, 1287–1288 (2002).
5. Kemp, A. C. et al. Climate related sea-level variations over the past two millennia. *Proc. Natl Acad. Sci. USA* **108**, 11017–11022 (2011).
6. Masson-Delmotte, V. et al. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. et al.) 383–464 (Cambridge Univ. Press, 2013).
7. Loutre, M. F. & Berger, A. Marine Isotope Stage 11 as an analogue for the present interglacial. *Global Planet. Change* **36**, 209–217 (2003).
8. Archer, D. & Ganopolski, A. A movable trigger: fossil fuel CO₂ and the onset of the next glaciation. *Geochim. Geophys. Res.* **6**, Q05003 (2005).
9. Paillard, D. What drives the Ice Age cycle? *Science* **313**, 455–456 (2006).
10. Petit, J. R. et al. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
11. Augustin, L. et al. Eight glacial cycles from an Antarctic ice core. *Nature* **429**, 623–628 (2004).
12. Lambeck, K. et al. in *Understanding Sea-Level Rise and Variability* (eds Church, J. A., Woodworth, P. L., Aarup, T. & Wilson, W. S.) 61–121 (Wiley-Blackwell, 2010).
13. Tzedakis, P. C., Channell, J. E. T., Hodell, D. A., Skinner, L. C. & Kleiven, H. F. Determining the natural length of the current interglacial. *Nature Geosci.* **5**, 138–141 (2012).
14. Petoukhov, V. et al. CLIMBER-2: a climate system model of intermediate complexity. Part I: model description and performance for present climate. *Clim. Dyn.* **16**, 1–17 (2000).
15. Greve, R. A continuum-mechanical formulation for shallow polythermal ice sheets. *Phil. Trans. R. Soc. A* **355**, 921–974 (1997).
16. Ganopolski, A. & Calov, R. The role of orbital forcing, carbon dioxide and regolith in 100 kyr cycles. *Clim. Past* **7**, 1415–1425 (2011).
17. Weertman, J. Milankovitch solar radiation variations and ice age ice sheet sizes. *Nature* **261**, 17–20 (1976).
18. Coehlin, A.-S., Mysak, L. A. & Wang, Z. Simulation of long-term future climate changes with the green McGill paleoclimate model: the next glacial inception. *Clim. Change* **79**, 381–401 (2006).
19. Calov, R. & Ganopolski, A. Multistability and hysteresis in the climate-cryosphere system under orbital forcing. *Geophys. Res. Lett.* **32**, L21717 (2005).
20. Abe-Ouchi, A. et al. Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume. *Nature* **500**, 190–193 (2013).
21. Ruddiman, W. F. The anthropogenic greenhouse era began thousands of years ago. *Clim. Change* **61**, 261–293 (2003).
22. Stocker, B. D., Strassmann, K. & Joos, F. Sensitivity of Holocene atmospheric CO₂ and the modern carbon budget to early human land use: analyses with a process-based model. *Biogeosciences* **8**, 69–88 (2011).
23. Ciais, P. et al. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. et al.) 465–570 (Cambridge Univ. Press, 2013).
24. Brovkin, V., Ganopolski, A., Archer, D. & Munhoven, G. Glacial CO₂ cycle as a succession of key physical and biogeochemical processes. *Clim. Past* **8**, 251–264 (2012).
25. Lisiecki, L. E. & Raymo, M. E. A. Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanography* **20**, PA1003 (2005).

Author Contributions A.G. and H.J.S. designed the paper. A.G. developed the methodology and performed the research (with contributions from R.W.). A.G., R.W. and H.J.S. interpreted the results and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.G. (ganopolski@pik-potsdam.de).

METHODS

Model. For this study we used the Earth system model of intermediate complexity CLIMBER-2 (ref. 14), which incorporates the three-dimensional thermo-mechanical ice sheet model SICOPOLIS¹⁵. SICOPOLIS is applied only to the Northern Hemisphere with a spatial resolution of $1.5^\circ \times 0.75^\circ$ and is fully interactively coupled to a low-resolution climate component. Key model parameters were selected to reproduce the magnitude of global ice volume variations during the last glacial cycle. Using the same set of model parameters, the model was then successfully applied to simulate the last eight glacial cycles¹⁶. The model also includes a global carbon cycle component that has been used to simulate CO₂ variations during the last glacial cycle²⁴. In the present study we use CLIMBER-2 in both configurations: (1) with interactive carbon cycle and prescribed modern ice sheets and (2) with interactive ice sheets and prescribed CO₂ concentration.

Model ensemble. Although the glacial cycles simulated with the standard version of CLIMBER-2 are in good agreement with palaeoclimate data¹⁶ and satisfy our palaeoclimate constraints it is possible that other model versions with slightly different model parameters are also compatible with palaeoclimate constraints. To assess how tightly palaeoclimate data constrain model parameters and to assess one of the sources of possible uncertainties in the CO₂–insolation relationship we created an ensemble of twenty CLIMBER-2 realizations by perturbing one of the model parameters: namely, parameter C_1 , which controls the effective cloud height in the atmospheric model component¹⁴. The advantage of this parameter is that it affects the equilibrium surface air temperature in a rather uniform manner, similar to changes in the CO₂ concentration. However, a number of other model parameters can be used for the same purpose. By varying this parameter by $\pm 5\%$ we created a set of model realizations that differ by only $\pm 1^\circ\text{C}$ in their global mean temperatures, a range that is comparable to the inter-model differences of CMIP-5 climate models²⁶.

From the twenty-member ensemble we then selected those simulations that are in agreement with observational constraints provided by glacial inceptions at the end of MIS11 and MIS5, and the lack of glacial inception in the recent past. To select the model realizations compatible with the empirical constraints we performed a set of simulations for MIS19, MIS11 and the current interglacial using the entire ensemble of the CLIMBER-2 model realizations. For the current interglacial, the model was run for 60,000 years into the future, starting 10,000 years before present from an equilibrium state. The MIS11 and MIS19 simulations begin 10,000 years before the respective insolation minima (410 kyr BP and 790 kyr BP, respectively), which triggered glacial inceptions. In all experiments the CO₂ concentration is fixed to the typical value for the given interglacial, that is, 280 p.p.m. for MIS11 and MIS1, and 240 p.p.m. for MIS19. These numbers are representative of the end of interglacials for MIS19, MIS11 and the late Holocene. Only four model realizations (including the standard one) were found to be consistent with all three palaeoclimate constraints.

Future simulations. Simulations of climate evolution over the next 100,000 years were performed in two steps. First, we computed the CO₂ concentration using the CLIMBER-2 model in the same configuration as in ref. 24. As initial conditions we used the state of the Earth system obtained at the end of the last glacial cycle simulation. (Note that the model simulates the present-day CO₂ concentration as close to the pre-industrial 280 p.p.m. level without any anthropogenic emissions.) Then the model was run for 100,000 years with the fully interactive carbon cycle forced by anthropogenic emissions and orbital parameters. Present-day ice sheets were prescribed during the entire runs. In the run without anthropogenic emissions, CO₂ gradually declines during the next 100,000 years owing to a small imbalance between volcanic outgassing and weathering. Assuming zero anthropogenic emissions beyond the next several centuries, the long-term CO₂ concentration depends only on the total cumulative carbon emissions. Therefore we used very simplistic emission scenarios which correspond to cumulative carbon emissions of 500 Gt C, 1,000 Gt C and 1,500 Gt C. For each emission scenario we performed one 100,000-year-long experiment with the standard version of the model. We then ran each of the four valid CLIMBER-2 model versions with the interactive ice sheets forced by the different CO₂ scenarios.

Insolation–CO₂ relationship. To find the relationship between insolation and CO₂ concentration we performed 20 simulations with 10 different orbital configurations and the coldest and warmest of the four valid model versions. Note that the difference in global surface air temperatures between these two model versions does not exceed 0.5°C under interglacial conditions. We chose orbital configurations that correspond to present-day conditions and eight previous minima of summer insolation, seven of which correspond to well defined glacial inceptions during the past 800 kyr. The last orbital configuration at 480 kyr BP, which corresponds to one of the shallowest minima in summer insolation, was chosen to span a broader range of summer insolation. All simulations started from the same pre-industrial equilibrium states both for the climate components

and ice sheets. Initial CO₂ concentration was set to 500 p.p.m. for the coldest three orbits and 400 p.p.m. for the other five. The orbital parameters were kept constant during each experiment and CO₂ concentration decreased linearly with time at the rate of 1 p.p.m. per 1,000 climate model years (see Extended Data Fig. 2). To be as close to the quasi-equilibrium for the ice sheets as possible, we used an acceleration technique with asynchronous coupling between climate and ice sheets, which effectively increases the length of integration for the ice sheet by a factor of 10. This implies that for the ice sheet component the rate of CO₂ change was only 1 p.p.m. per 10,000 years. Each experiment lasted 200,000 climate model years or 2 million ice-sheet model years.

We performed several additional experiments with a slower rate of change and found that this rate is sufficient to detect ‘glacial inception’ with an accuracy of about 5 p.p.m. in the CO₂ space. As a criterion for glacial inception we consider the growth of Northern Hemisphere ice sheets (excluding Greenland) of more than 5 m in sea-level equivalent (Extended Data Fig. 2). With the early model version it was shown¹⁹ that for a given CO₂ concentration the onset of glacial inception to the first approximation depends only on the maximum summer insolation at 65°N . This is also supported by Fig. 3a. We used least-squares methods to fit the data to the logarithmic curve. The choice of a logarithmic curve is natural because the radiative forcing of CO₂ is proportional to the logarithm of CO₂ concentration.

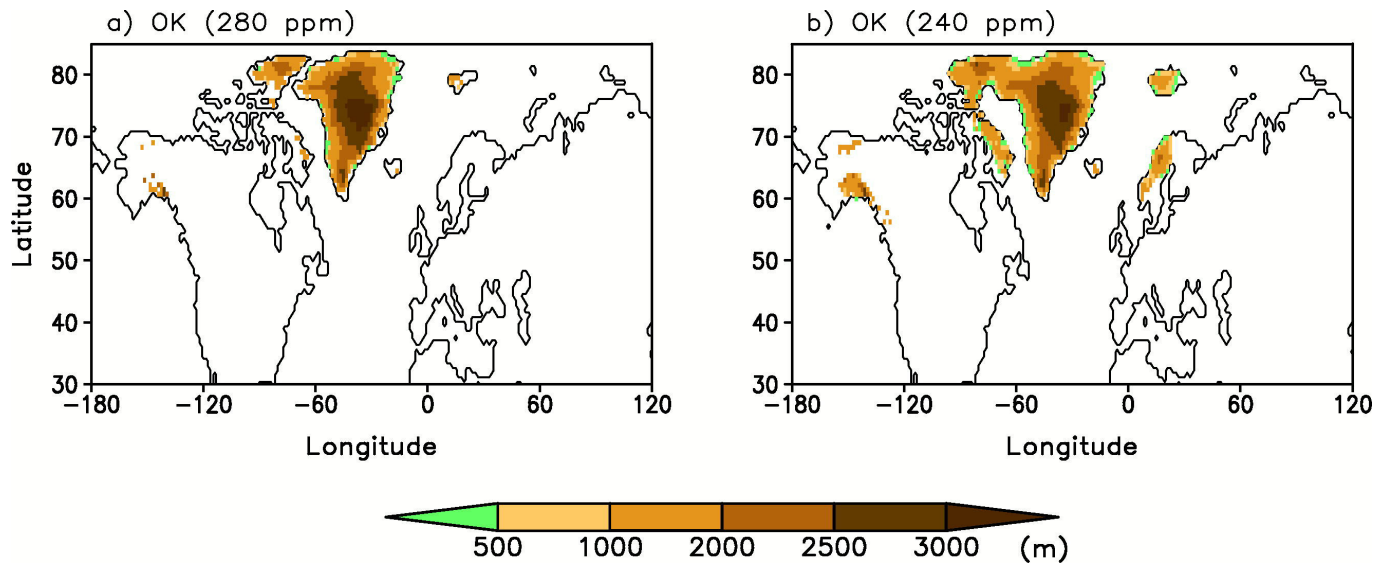
The shaded area in Fig. 3a represents the uncertainty range of critical insolation $\pm 4\text{ W m}^{-2}$ (one standard deviation). In the vicinity of pre-industrial CO₂ concentration, this uncertainty in insolation corresponds to the uncertainty in the CO₂ threshold value of $\pm 20\text{ p.p.m.}$ The absolute uncertainties are lower for low CO₂ and higher for high CO₂ concentrations. This total uncertainty originates from three sources: (1) uncertainties resulting from the existence of several model versions, all of which are consistent with our palaeoclimate constraints, (2) the fact that the critical insolation depends not only on total summer insolation, but, to a lesser degree, also on individual contributions from precession and obliquity¹⁹, and (3) errors related to the methods of tracing the threshold CO₂ values. Although past glacial inceptions provide a tight constraint on a range of model parameters (Fig. 3), it is necessary to note that a single-model ensemble obtained by perturbation of model parameters may considerably underestimate the range of uncertainties compared to multi-model ensembles, owing to the lack of structural uncertainties. More objective estimates of uncertainties in the insolation–CO₂ relationship would therefore only be possible when a number of different Earth system models are available for this type of study.

We found that our new insolation–CO₂ relationship differs from that in ref. 8, primarily in that the new insolation–CO₂ graph is concave, unlike the convex shape in ref. 8. This difference can be explained by the much faster rates of changes of the forcing used in ref. 8, which do not allow us to detect bifurcation points with sufficient accuracy. The new methodology and greatly increased computer performance allow us now to determine the critical CO₂ value much more accurately to obtain a physically sound insolation–CO₂ relationship.

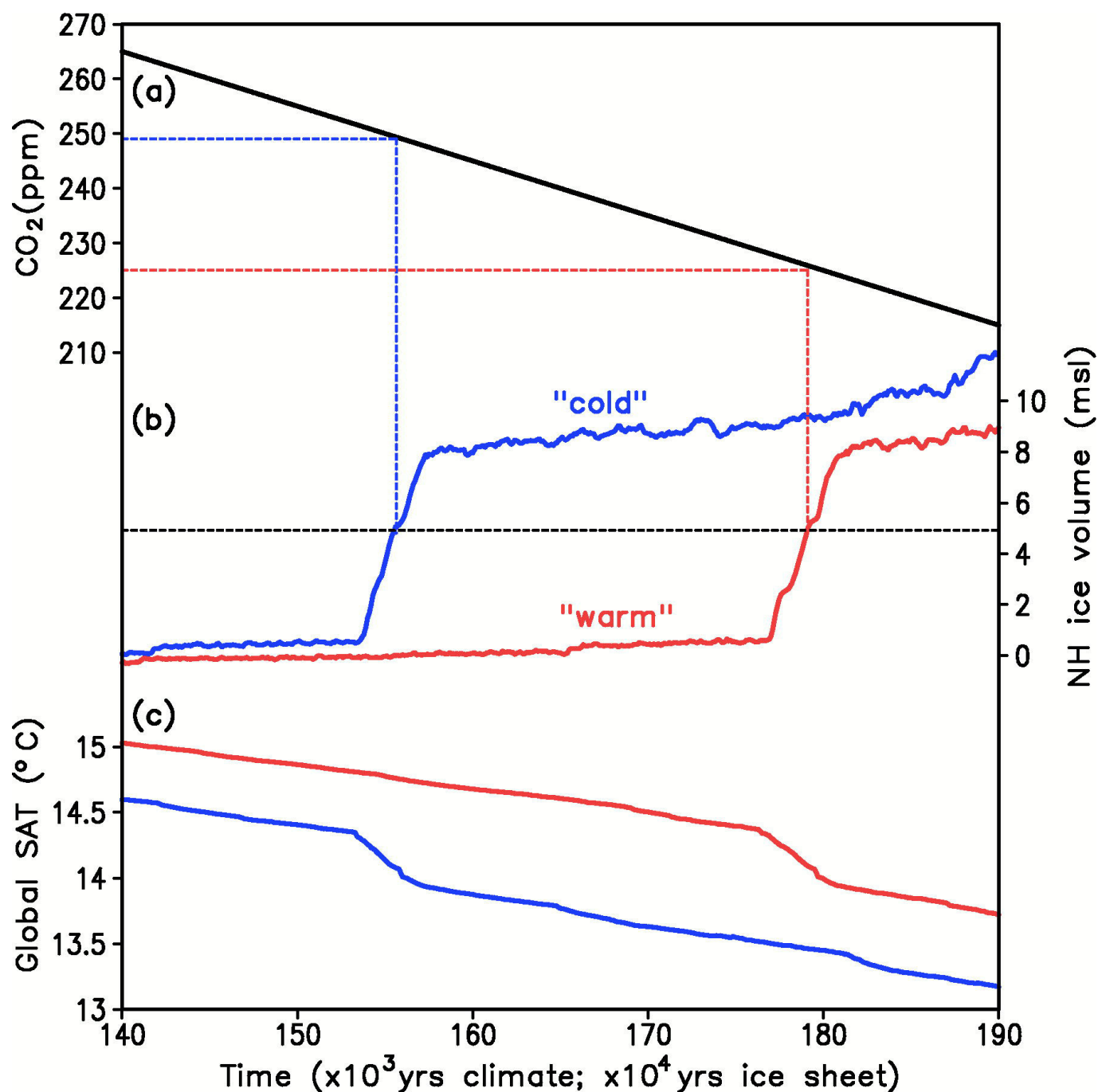
Timing of past and future glacial inceptions. Using the insolation–CO₂ relationship and orbital parameters²⁷ we computed the value of the critical CO₂ concentration every 1,000 years (Fig. 3c and Extended Data Fig. 3). If the critical CO₂ concentration is above its actual value and the system is in an interglacial state then glacial inception should begin. When the system is already in a glacial state the critical relationship is not applicable. This is because the strong nonlinearity and extremely long response time of the climate–cryosphere system mean that its dynamics depend not only on the instantaneous forcing but also on its past evolution. Therefore our criterion for glacial inception can only be applied to interglacial states, that is, under conditions similar to the Holocene with respect to both global ice volume and temperatures. Some interglacials, such as the Holocene (MIS1), Eem (MIS5e), MIS9 and MIS11, are clearly defined in the various palaeoclimate records (Extended Data Fig. 3). The situation is less clear for MIS7 and MIS15, where several interglacials per one glacial cycle can be identified²⁸. For these two cases we show in Fig. 3 two different glacial inceptions.

Code availability. Code for the ice sheet model SICOPOLIS can be accessed at <http://www.sicopolis.net>. Code for the Earth system model CLIMBER-2 model is not available owing to a lack of comprehensive technical description.

26. Mauritsen, T. *et al.* Tuning the climate of a global model. *J. Adv. Model. Earth Syst.* **4**, M00A01 (2012).
27. Laskar, J. *et al.* A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.* **428**, 261–285 (2004).
28. Yin, Q. & Berger, A. Individual contribution of insolation and CO₂ to the interglacial climates of the past 800,000 years. *Clim. Dyn.* **38**, 709–724 (2012).
29. Elderfield, H. *et al.* Evolution of ocean temperature and ice volume through the mid-Pleistocene climate transition. *Science* **337**, 704–709 (2012).
30. Grant, K. M. *et al.* Sea-level variability over five glacial cycles. *Nature Commun.* **5**, 5076 (2014).

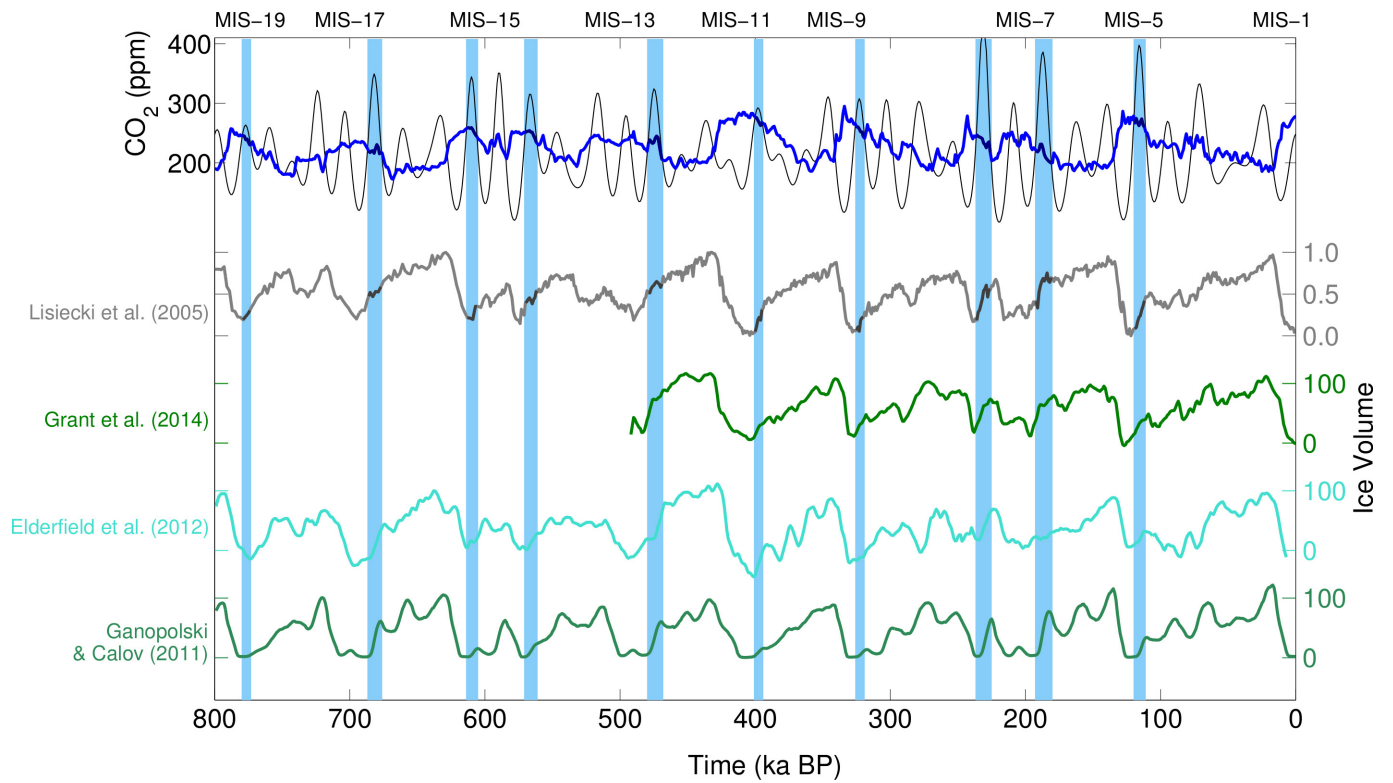


Extended Data Figure 1 | Ice sheet at 0 kyr BP. The extension and elevation of simulated Northern Hemisphere ice sheets at the time corresponding to present-day (0 kyr BP, 'OK') insolation are shown for constant CO₂ concentrations of 280 p.p.m. (a) and 240 p.p.m. (b). Experiments were performed with the coldest of the accepted model versions.



Extended Data Figure 2 | Detection of the critical CO_2 value. For the orbital configuration corresponding to 777 kyr BP we show the prescribed CO_2 concentration (a), the simulated Northern Hemisphere (NH) ice volume in metres of sea-level equivalent (msl; b), and the simulated global mean surface air temperature (SAT) (c). Blue lines correspond to

the coldest of accepted model versions and red lines to the warmest. The figure shows only the part of 200,000-year-long simulations for which glacial inceptions are simulated by both model versions. We note that the simulation time is ten times larger for the ice-sheet component than for the climate component (see bottom axis).



Extended Data Figure 3 | Past glacial inceptions. Past glacial inceptions, detected using the critical insolation–CO₂ relation, are shown in comparison to three different reconstructions of ice-volume variations over the past 800 kyr (refs 25, 29 and 30) and results of model simulations¹⁶.

Plant functional traits have globally consistent effects on competition

Georges Kunstler^{1,2,3}, Daniel Falster³, David A. Coomes⁴, Francis Hui⁵, Robert M. Kooyman^{3,6}, Daniel C. Laughlin⁷, Lourens Poorter⁸, Mark Vanderwel⁹, Ghislain Vieilledent¹⁰, S. Joseph Wright¹¹, Masahiro Aiba¹², Christopher Baraloto^{13,14}, John Caspersen^{15,16}, J. Hans C. Cornelissen¹⁷, Sylvie Gourlet-Fleury¹⁰, Marc Hanewinkel^{18,19}, Bruno Herault²⁰, Jens Kattge^{21,22}, Hiroko Kurokawa^{12†}, Yusuke Onoda²³, Josep Peñuelas^{24,25}, Hendrik Poorter²⁶, Maria Uriarte²⁷, Sarah Richardson²⁸, Paloma Ruiz-Benito^{29,30}, I-Fang Sun³¹, Göran Ståhl³², Nathan G. Swenson³³, Jill Thompson^{34,35}, Bertil Westerlund³², Christian Wirth^{22,36}, Miguel A. Zavala³⁰, Hongcheng Zeng¹⁵, Jess K. Zimmerman³⁵, Niklaus E. Zimmermann¹⁶ & Mark Westoby³

Phenotypic traits and their associated trade-offs have been shown to have globally consistent effects on individual plant physiological functions^{1–3}, but how these effects scale up to influence competition, a key driver of community assembly in terrestrial vegetation, has remained unclear⁴. Here we use growth data from more than 3 million trees in over 140,000 plots across the world to show how three key functional traits—wood density, specific leaf area and maximum height—consistently influence competitive interactions. Fast maximum growth of a species was correlated negatively with its wood density in all biomes, and positively with its specific leaf area in most biomes. Low wood density was also correlated with a low ability to tolerate competition and a low competitive effect on neighbours, while high specific leaf area was correlated with a low competitive effect. Thus, traits generate trade-offs between performance with competition versus performance without competition, a fundamental ingredient in the classical hypothesis that the coexistence of plant species is enabled via differentiation in their successional strategies⁵. Competition within species was stronger than between species, but an increase in trait dissimilarity between species had little influence in weakening competition. No benefit of dissimilarity was detected for specific leaf area or wood density, and only a weak benefit for maximum height. Our trait-based approach to modelling competition makes generalization possible across the forest ecosystems of the world and their highly diverse species composition.

Phenotypic traits are considered fundamental drivers of community assembly and thus species diversity^{1,6}. The effects of traits on individual plant physiologies and functions are increasingly understood, and have been shown to be underpinned by well-known and globally consistent trade-offs^{1–3}. For instance, traits such as wood density and specific leaf area capture trade-offs between the construction cost and longevity or

strength of wood and leaf tissues^{2,3}. By contrast, we still have a limited understanding of how such trait-based trade-offs translate into competitive interactions between species, particularly for long-lived organisms such as trees. Competition is a key filter through which ecological and evolutionary success is determined⁴. A long-standing hypothesis is that the intensity of competition decreases as two species diverge in trait values⁷ (trait dissimilarity). The few studies^{8–13} that have explored links between traits and competition have shown that linkages were more complex than this, as particular trait values may also confer competitive advantage independently from trait dissimilarity^{9,13,14}. This distinction is fundamental for species coexistence and the local mixture of traits. If neighbourhood competition is driven mainly by trait dissimilarity, this will favour a wide spread of trait values at a local scale. By contrast, if neighbourhood interactions are mainly driven by the competitive advantage associated with particular trait values, those trait values should be strongly selected at the local scale, with coexistence operating at larger spatial or temporal scales^{6,13}. Empirical investigations have been limited so far to a few particular locations, restricting our ability to find general mechanisms that link traits and competition in the main vegetation types of the world.

Here we quantify the links between traits and competition, measured as the influence of neighbouring trees on growth of a focal tree. Our framework is novel in two important ways: first, competition is analysed at an unprecedented scale covering all the major forest biomes on Earth (Fig. 1a and Extended Data Fig. 1), and second, the influence of traits on competition is partitioned among four fundamental mechanisms (Fig. 1b, c) as follows. A competitive advantage for trees with some trait values compared to others can arise by: (1) permitting faster maximum growth in the absence of competition¹⁵; (2) exerting a stronger competitive effect^{16,17}, meaning that competitor species possessing those

¹Irstea, UR EMGR, 2 rue de la Papeterie BP-76, F-38402, St-Martin-d'Hères, France. ²Univ. Grenoble Alpes, F-38402 Grenoble, France. ³Department of Biological Sciences, Macquarie University, New South Wales 2109, Australia. ⁴Forest Ecology and Conservation Group, Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK. ⁵Mathematical Sciences Institute, The Australian National University, Canberra 0200, Australia. ⁶National Herbarium of New South Wales, Royal Botanic Gardens and Domain Trust, Sydney 2000, New South Wales, Australia. ⁷Environmental Research Institute, School of Science, University of Waikato, Hamilton 3240, New Zealand. ⁸Forest Ecology and Forest Management Group, Wageningen University, 6708 PB Wageningen, The Netherlands. ⁹Department of Biology, University of Regina, 3737 Wascana Pkwy, Regina SK S4S 0A2, Canada. ¹⁰Cirad, UPR BSEF, F-34398 Montpellier, France. ¹¹Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Republic of Panama. ¹²Graduate School of Life Sciences, Tohoku University, Sendai 980-8578, Japan. ¹³INRA, UMR Ecologie des Forêts de Guyane, BP 709, 97387 Kourou Cedex, France. ¹⁴International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Miami, Florida 33199, USA. ¹⁵Faculty of Forestry, University of Toronto, 33 Willcocks Street, Toronto, Ontario M5S 3B3, Canada. ¹⁶Swiss Federal Research Institute WSL, Landscape Dynamics Unit, CH-8903 Birmensdorf, Switzerland. ¹⁷Systems Ecology, Department of Ecological Science, Vrije Universiteit, Amsterdam 1081 HV, The Netherlands. ¹⁸Swiss Federal Research Institute WSL, Forest Resources and Management Unit, CH-8903 Birmensdorf, Switzerland. ¹⁹University of Freiburg, Chair of Forestry Economics and Planning, D-79106 Freiburg, Germany. ²⁰Cirad, UMR Ecologie des Forêts de Guyane, Campus Agronomique, BP 701, 97387 Kourou, France. ²¹Max Planck Institute for Biogeochemistry, Hans Knöll Str. 10, 07745 Jena, Germany. ²²German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Deutscher Platz 5e 04103 Leipzig, Germany. ²³Graduate School of Agriculture, Kyoto University, Kyoto, 606-8502 Japan. ²⁴CSIC, Global Ecology Unit CREAL-CSIC-UAB, Cerdanyola del Vallès 08193, Catalonia, Spain. ²⁵CREAF, Cerdanyola del Vallès, 08193 Barcelona, Catalonia, Spain. ²⁶Plant Sciences (IBG-2), Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany. ²⁷Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, New York 10027, USA. ²⁸Landcare Research, PO Box 40, Lincoln 7640, New Zealand. ²⁹Biological and Environmental Sciences, School of Natural Sciences, University of Stirling, Stirling FK9 4LA, UK. ³⁰Forest Ecology and Restoration Group, Department of Life Sciences, Science Building, University of Alcalá, Campus Universitario, 28805 Alcalá de Henares (Madrid), Spain. ³¹Department of Natural Resources and Environmental Studies, National Dong Hwa University, Hualien 97401, Taiwan. ³²Department of Forest Resource Management, Swedish University of Agricultural Sciences (SLU), Skogsmarksgränd, 901 83 Umeå, Sweden. ³³Department of Biology, University of Maryland, College Park, Maryland 20742, USA. ³⁴Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian EH26 0QB, UK. ³⁵Department of Environmental Sciences, University of Puerto Rico, Río Piedras Campus PO Box 70377 San Juan, Puerto Rico 00936-8377, USA. ³⁶Institute for Systematic, Botany and Functional Biodiversity, University of Leipzig, Johannisallee 21 04103 Leipzig, Germany. [†]Present address: Forestry and Forest Products Research Institute, Tsukuba 305-8687, Japan.

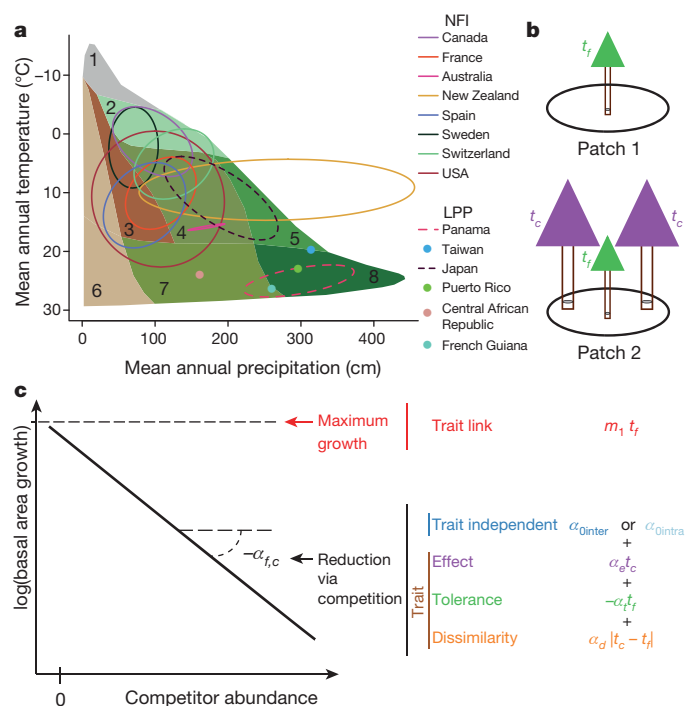


Figure 1 | Assessing competitive interactions at global scale.

a, Precipitation–temperature space occupied by each data set (LPP, large permanent plots data; NFI, national forest inventories data). For data with multiple plots, the range of climatic condition is represented by an ellipse covering 98% of the plots. Biomes are: 1, tundra; 2, taiga; 3, Mediterranean; 4, temperate forest; 5, temperate rainforest; 6, desert; 7, tropical seasonal forest; and 8, tropical rainforest (see Methods for details). **b**, Sampled patches vary in the abundance of competitors from species *c* around individuals of focal species *f*. **c**, We modelled how trait values of the focal tree (t_f), and the abundance (measured as the sum of their basal areas) and traits values of competitor species (t_c) influenced basal area growth of the focal tree. Species maximum growth (red) was influenced by trait of the focal tree ($m_0 + m_1 t_f$, with m_0 maximum growth independent of the trait). Reduction in growth per unit basal area of competitors ($-\alpha_{f,c}$, black) was modelled as the sum of growth reduction independent of the trait (blue) by conspecific ($\alpha_{0\text{intra}}$) and heterospecific ($\alpha_{0\text{inter}}$) competitors, the effect of competitor traits (t_c) on their competitive effect (α_e), the effect of the focal tree's traits (t_f) on its tolerance of competition (α_t), and the effect of trait dissimilarity between the focal tree and its competitors ($|t_c - t_f|$) on competition (α_d). The parameters m_0 , m_1 , $\alpha_{0\text{intra}}$, $\alpha_{0\text{inter}}$, α_e , α_t and α_d are fitted from data using a maximum likelihood method.

traits suppress more strongly the growth of their neighbours; or (3) permitting a better tolerance of competition (described as a competitive 'response' in ref. 16), meaning that the growth of species possessing those traits is less affected by competition from neighbours. Finally, (4) competition can promote trait diversification, if increasing trait dissimilarity between species reduces interspecific competition compared to intraspecific competition⁷. Here we show how these four mechanisms are connected to three key traits that describe plant strategies worldwide^{1–3}. These traits are wood density (an indicator of a trade-off in stems between growth and strength), specific leaf area (SLA; an indicator of a trade-off in leaves between cheap construction cost and leaf longevity), and maximum height (an indicator of a trade-off between sustained access to light and early reproduction). We analyse the basal area growth (annual increase in the cross-section area of a tree trunk at 1.3 m height) of more than 3 million trees from over 2,500 species, across all major forested biomes of the Earth (Fig. 1). Species mean trait values were extracted from local data bases and the global TRY data base¹⁸ (see Methods). We analysed how the basal area growth of each individual tree was reduced by the abundance of competitors in its local neighbourhood¹⁹ (measured as the sum of basal areas of competitors in $\text{m}^2 \text{ha}^{-1}$), accounting for traits of both

the focal tree and its competitors. This analysis allowed effect sizes to be estimated for each of the four mechanisms outlined earlier (Fig. 1c).

Across all biomes, the strongest driver of individual growth was the total abundance of neighbours, irrespective of their traits (parameters $\alpha_{0\text{intra}}$ and $\alpha_{0\text{inter}}$ in Fig. 2). Values were strongly positive, indicating that neighbours had competitive rather than facilitative effects. The main effects of traits were that some trait values led to a competitive advantage compared to others through two main mechanisms. First, traits of the focal species had influences on its maximum growth—that is, in the absence of competition—(parameter m_1 in Fig. 2 and Extended Data Table 1). The fastest growing species had low wood density and high SLA values, although the confidence interval on the trait effect intercepted zero in two out of five biomes for SLA (Fig. 2). This is in agreement with previous studies^{15,20} of adult trees reporting a strong link between maximum growth and wood density but a weaker link for SLA. Second, some trait values were associated with species having stronger competitive effects, or better tolerance of competition (Fig. 2 and Extended Data Table 1). High wood density was correlated with better tolerance of competition from neighbours and with a stronger competitive effect on neighbours, whereas low SLA was correlated only with a stronger competitive effect. This agrees with studies reporting that high wood density species are more shade-tolerant¹⁵ and have deeper and wider crowns^{21,22}, hence potentially higher light interception (further detail in Supplementary Discussion). The shorter leaf lifespan associated with high SLA results in lower leaf mass fraction²³. The low competitive effect associated with high SLA species could thus result from a lower light interception, but few data are available on this link²³. Maximum height was weakly negatively correlated with tolerance to competition in three out of five biomes, supporting the idea that sub-canopy trees are more shade-tolerant²¹. We found, however, no correlation between maximum height and competitive effect. The current height of an individual does have an influence on light interception, a key process in competition¹³. But maximum height of a species reflects its long-term strategy, and would possibly have stronger effects on long-term population level competition outcomes than it did on short-term basal area growth²⁴.

After separating trait-independent differences between intraspecific versus interspecific competition, trait dissimilarity had little effect on competition between species (Fig. 2). Only dissimilarity in maximum height between focal and neighbouring species led to a weak, but consistent, decrease in competitive suppression of tree growth (Fig. 2). Mechanisms explaining this effect are poorly understood, but could possibly result from complementary crown architectures^{25,26}. The average differences in strength of interspecific versus intraspecific competition between two species—a key indicator of processes that could stabilize coexistence—were thus only weakly related to trait dissimilarity (Extended Data Fig. 2). Trait dissimilarity effects are widely considered to be a key mechanism by which traits affect competition¹³, but our analysis shows at global scale that trait dissimilarity effects are weak or absent. It remains unclear why the trait-independent competitive effects are higher within species than between species. Higher loads of shared specialized pathogens²⁷ could plausibly contribute. Other traits or combinations of traits (see ref. 12) may show stronger trait dissimilarity effects, but we currently lack the trait data to capture such effects.

Analyses allowing for different effects among biomes did not show any particular biome behaving consistently differently from the others (Fig. 2). This lack of context dependence in trait effects may seem surprising, but reinforces the idea that competition for light is important in most forests, and this may explain why we find consistency across such diverse forest types (see Supplementary Discussion for further details).

Our global study supports the hypothesis that trait values favouring high tolerance of competition or high competitive effects also render species slow-growing in the absence of competition across all forested biomes (Fig. 3). This trait-based trade-off is a key ingredient in the classical model of successional coexistence in forests, in which fast-growing species are more abundant in early successional

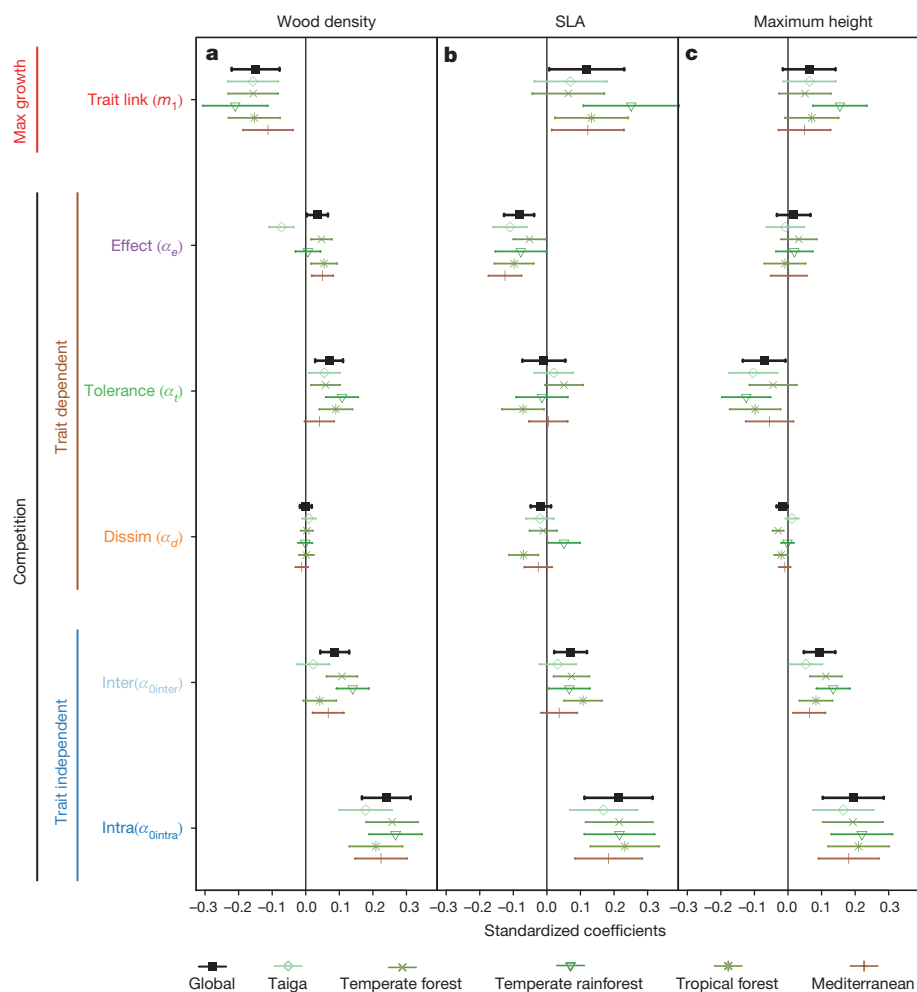


Figure 2 | Trait-dependent and trait-independent effects on maximum growth and competition across the globe, and their variation among biomes. a–c, Standardized regression coefficients for growth models, fitted separately for wood density (a), SLA (b) and maximum height (c) (points denote average estimates, lines denote 95% confidence intervals). Black points and lines represent global estimates, and coloured points and lines represent the biome level estimates. The parameter estimates represent: the effect of focal tree's trait value on maximum growth m_1 , the effect of competitor trait values on their competitive effect α_e (positive values indicate that higher trait values lead to a stronger reduction in growth of the focal tree), the effect of the focal tree's trait value on its tolerance of competition α_t (positive values indicate that greater trait values result in greater tolerance of competition), the effect on competition of trait dissimilarity between the focal tree and its competitors α_d (negative values indicate that higher trait dissimilarity leads to a lower reduction of the growth of the focal tree), and the trait-independent competitive effect of conspecific (α_{0intra}) and heterospecific (α_{0inter}) competitors. Tropical rainforest and tropical seasonal forest were merged together as tropical forest, tundra was merged with taiga, and desert was not included as too few plots were available (see Fig. 1a for biomes definitions).

stages where competitors are absent or rare, and are later replaced by slow-growing species in late successional stages where competitors become more abundant⁵. Human or natural disturbances are conspicuous in all of the forests analysed, hence successional dynamics are

likely to be present in all these sites (see Supplementary Information). This trade-off was strongest for wood density, with high wood density associated with a slow potential growth rate but a high tolerance to competition and a strong competitive effect (Fig. 3). A similar pattern

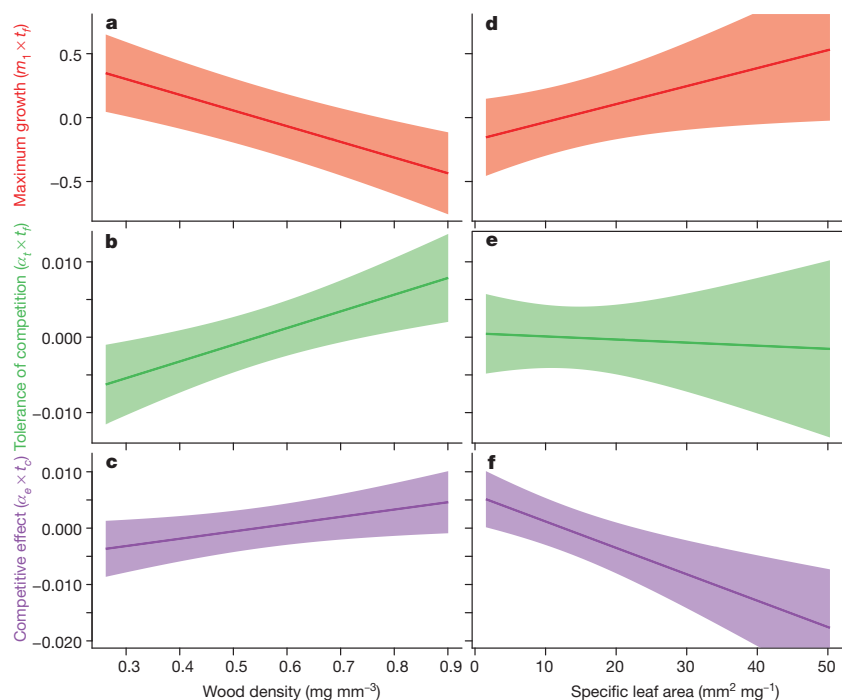


Figure 3 | Variation of maximum growth, competitive effects and competitive tolerance with wood density and SLA predicted by global traits models. a–f, Variation of maximum growth ($m_1 \times t_j$) (a, d), tolerance of competition ($\alpha_t \times t_j$) (b, e) and competitive effect ($\alpha_e \times t_j$) (c, f) parameters with wood density (a–c) and SLA (d–f). The shaded area represents the 95% confidence interval of the prediction (including uncertainty associated with α_0 or m_0).

was present, although less clear, for SLA. High SLA was correlated with a low competitive effect but fast maximum growth (confidence intervals not spanning zero in three biomes, Figs 2 and 3, see Extended Data Fig. 3 for maximum height). Given that the long-term outcomes of competition at the population level may be more influenced by tolerance of competition than by the competitive effect¹⁶, SLA might be less influential in succession.

Coordination between trait values conferring a strong competitive effect and trait values conferring a high tolerance of competition has been widely expected^{9,16}, but rarely documented^{16,28}. Only wood density showed such coordination, as it was correlated with both competitive effect and tolerance of competition in the same direction (Fig. 2).

The globally consistent links that we report here between traits and competition have considerable promise for predicting species interactions governing forest communities across different forest biomes and continents of the globe. Our analysis demonstrates that trait dissimilarity is not the major determinant of local-scale competitive effects on tree growth, at least for these three traits. By contrast, the trait-based trade-off in performance with competition versus without competition, reported here, could promote the coexistence of species with diverse traits, provided disturbances create a mosaic of successional stages. A challenge for the future is to move beyond growth to analyse all key demographic rates and life-history stages, and analyse how traits influence competitive outcomes and stable coexistence at the population level.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 August; accepted 23 November 2015.

Published online 23 December 2015.

- Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. a. & Wright, I. J. Plant Ecological Strategies: some leading dimensions of variation between species. *Annu. Rev. Ecol. Syst.* **33**, 125–159 (2002).
- Wright, I. J. *et al.* The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).
- Chave, J. *et al.* Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12**, 351–366 (2009).
- Keddy, P. A. *Competition* (Springer Netherlands, 1989).
- Rees, M., Condit, R., Crawley, M., Pacala, S. W. & Tilman, D. Long-term studies of vegetation dynamics. *Science* **293**, 650–655 (2001).
- Adler, P. B., Fajardo, A., Kleinhesselink, A. R. & Kraft, N. J. B. Trait-based tests of coexistence mechanisms. *Ecol. Lett.* **16**, 1294–1306 (2013).
- MacArthur, R. & Levins, R. The limiting similarity, convergence, and divergence of coexisting species. *Am. Nat.* **101**, 377–385 (1967).
- Uriarte, M. *et al.* Trait similarity, shared ancestry and the structure of neighbourhood interactions in a subtropical wet forest: implications for community assembly. *Ecol. Lett.* **13**, 1503–1514 (2010).
- Kunstler, G. *et al.* Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecol. Lett.* **15**, 831–840 (2012).
- HilleRisLambers, J., Adler, P., Harpole, W., Levine, J. & Mayfield, M. Rethinking community assembly through the lens of coexistence theory. *Annu. Rev. Ecol. Syst.* **43**, 227–248 (2012).
- Lasky, J. R., Uriarte, M., Boukili, V. K. & Chazdon, R. L. Trait-mediated assembly processes predict successional changes in community diversity of tropical forests. *Proc. Natl Acad. Sci. USA* **111**, 5616–5621 (2014).
- Kraft, N. J. B., Godoy, O. & Levine, J. M. Plant functional traits and the multidimensional nature of species coexistence. *Proc. Natl Acad. Sci. USA* **112**, 797–802 (2015).
- Mayfield, M. M. & Levine, J. M. Opposing effects of competitive exclusion on the phylogenetic structure of communities: phylogeny and coexistence. *Ecol. Lett.* **13**, 1085–1093 (2010).
- Kraft, N. J. B., Crutsinger, G. M., Forrester, E. J. & Emery, N. C. Functional trait differences and the outcome of community assembly: an experimental test with vernal pool annual plants. *Oikos* **123**, 1391–1399 (2014).
- Wright, S. J. *et al.* Functional traits and the growth-mortality trade-off in tropical trees. *Ecology* **91**, 3664–3674 (2010).
- Goldberg, D. E. Competitive ability: definitions, contingency and correlated traits. *Phil. Trans. R. Soc. Lond. B* **351**, 1377–1385 (1996).
- Gaudet, C. L. & Keddy, P. A. A comparative approach to predicting competitive ability from plant traits. *Nature* **334**, 242–243 (1988).
- Kattge, J. *et al.* TRY – a global database of plant traits. *Glob. Change Biol.* **17**, 2905–2935 (2011).
- Uriarte, M., Canham, C. D., Thompson, J. & Zimmerman, J. K. A neighborhood analysis of tree growth and survival in a hurricane-driven tropical forest. *Ecol. Monogr.* **74**, 591–614 (2004).
- Poorter, L. *et al.* Are functional traits good predictors of demographic rates? Evidence from five neotropical forests. *Ecology* **89**, 1908–1920 (2008).
- Poorter, L., Bongers, L. & Bongers, F. Architecture of 54 moist-forest tree species: traits, trade-offs, and functional groups. *Ecology* **87**, 1289–1301 (2006).
- Aiba, M. & Nakashizuka, T. Architectural differences associated with adult stature and wood density in 30 temperate tree species. *Funct. Ecol.* **23**, 265–273 (2009).
- Niinemets, Ü. A review of light interception in plant stands from leaf to canopy in different plant functional types and in species with varying shade tolerance. *Ecol. Res.* **25**, 693–714 (2010).
- Adams, T. P., Purves, D. W. & Pacala, S. W. Understanding height-structured competition in forests: is there an R* for light? *Proc. R. Soc. B* **274**, 3039–3047 (2007).
- Sapijanskas, J., Paquette, A., Potvin, C., Kunert, N. & Loreau, M. Tropical tree diversity enhances light capture through crown plasticity and spatial and temporal niche differences. *Ecology* **95**, 2479–2492 (2014).
- Jucker, T., Bouriaud, O. & Coomes, D. A. Crown plasticity enables trees to optimize canopy packing in mixed-species forests. *Funct. Ecol.* **29**, 1078–1086 (2015).
- Bagchi, R. *et al.* Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* **506**, 85–88 (2014).
- Wang, P., Stieglitz, T., Zhou, D. W. & Cahill, J. F. Jr. Are competitive effect and response two sides of the same coin, or fundamentally different? *Funct. Ecol.* **24**, 196–207 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are especially grateful to the researchers whose long-term commitment to establish and maintain forest plots and their associated databases made this study possible, and to those who granted us data access: forest inventories and permanent plots of New Zealand, Spain (MAGRAMA), France, Switzerland, Sweden, US and Canada (for the provinces of Quebec provided by the Ministère des Ressources Naturelles du Québec, Ontario provided by OnTAP's Growth and Yield Program of the Ontario Ministry of Natural Resources, Saskatchewan, Manitoba, New Brunswick, Newfoundland and Labrador), CTFs (BCI and LTER-Luquillo), Taiwan (Fushan), Cirad (Paracou with funding by CEBA, ANR-10-LABX-25-01), Cirad, MEFCP and ICRA (M'Baïki) and Japan. We thank MPI-BGC Jena, who host TRY, and the international funding networks supporting TRY (IGBP, DIVERSITAS, GLP, NERC, QUEST, FRB and GIS Climate). G.K. was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Program (Demo-Traits project, no. 299340). The working group that initiated this synthesis was supported by Macquarie University and by Australian Research Council through a fellowship to M.W.

Author Contributions G.K. and M.W. conceived the study, and with D.F. led a workshop with the participation of D.A.C., F.H., R.M.K., D.C.L., L.P., M.V., G.V. and S.J.W. G.K. wrote the manuscript with key inputs from all workshop participants and help from all authors. G.K., D.F. and F.H. wrote the computer code and processed the data. G.K. devised the main analytical approach and performed analyses with assistance from D.F. for the figures. G.K., D.A.C., D.F., F.H., R.M.K., D.C.L., M.V., G.V., S.J.W., M.A., C.B., J.C., J.H.C.C., S.G.-F., M.H., B.H., J.K., H.K., Y.O., J.P., H.P., M.U., S.R., P.R.-B., I.-F.S., G.S., N.G.S., J.T., B.W., C.W., M.A.Z., H.Z., J.K.Z. and N.E.Z. collected and processed the raw data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.K. (georges.kunstler@irstea.fr).

METHODS

Model and analysis. To examine the link between competition and traits, we used a neighbourhood modelling framework^{8,29} to model the growth of a focal tree of species (f) as a product of its maximum growth (determined by its traits and size) together with reductions due to competition from individuals growing in the local neighbourhood (see definition below). Specifically, we assumed a relationship of the form

$$G_{i,f,p,s,t} = G_{\max,f,p,s} D_{i,f,p,s,t}^{\gamma_f} \exp \left(\sum_{c=1}^{N_i} -\alpha_{f,c} B_{i,c,p,s} \right) \quad (1)$$

in which $G_{i,f,p,s,t}$ and $D_{i,f,p,s,t}$ are the annual basal area growth (G) and diameter (D) at breast height of individual (i), from species (f), plot or quadrat (p) (see below), data set (s) and census (t). $G_{\max,f,p,s}$ is the maximum basal area growth for species (f) on plot or quadrat (p), in data set (s); that is, in absence of competition, γ_f determines the rate at which growth changes with size for species (f), modelled with a normally distributed random effect of species ε_{γ_f} (as $\gamma_f = \gamma_0 + \varepsilon_{\gamma_f}$, in which $\varepsilon_{\gamma_f} \sim (0, \sigma_{\gamma})$)—a normal distribution of mean 0 and standard deviation σ_{γ}). $\alpha_{f,c}$ is the per unit basal area effect of individuals from species (c), on growth of an individual in species (f). $B_{i,c,p,s,t} = 0.25\pi \sum_{j \neq i} w_j D_{j,c,p,s,t}^2$ is the sum of basal area of all individuals competitor trees (j), of the species (c), within the local neighbourhood of the tree (i), in plot (p), data set (s), and census (t), where w_j is a constant based on neighbourhood size for tree (j), depending on the data set (see below). Note that $B_{i,c,p,s,t}$ include all trees of species (c) in the local neighbourhood except the tree (i), and N_i is the number of competitor species in the local neighbourhood of focal tree (i). Values of $\alpha_{f,c} > 0$ indicate competition, whereas $\alpha_{f,c} < 0$ indicates facilitation.

A log-transformation of equation (1) leads to a linearized model of the form

$$\log(G_{i,f,p,s,t}) = \log(G_{\max,f,p,s}) + \gamma_f \log(D_{i,f,p,s,t}) + \sum_{c=1}^{N_i} -\alpha_{f,c} B_{i,c,p,s,t} \quad (2)$$

To include the effects of traits on the parameters of the growth model, we build on previous studies that explored the role of traits for tree performances and tree competition^{8,9,11}. We modelled the effect of traits, one trait at a time. The effect of a focal species' trait value (t_f) on its maximum growth was included as:

$$\log(G_{\max,f,p,s}) = m_0 + m_1 t_f + m_2 \text{MAT} + m_3 \text{MAP} + \varepsilon_{G_{\max,f}} + \varepsilon_{G_{\max,p}} + \varepsilon_{G_{\max,s}} \quad (3)$$

Here, m_0 is the average maximum growth, m_1 gives the effect of the focal species trait, m_2 and m_3 the effects of mean annual temperature (MAT) and sum of annual precipitation (MAP), respectively, and $\varepsilon_{G_{\max,f}}$, $\varepsilon_{G_{\max,p}}$ and $\varepsilon_{G_{\max,s}}$ are normally distributed random effects for species (f), plot or quadrat (p) (see below), and data set (s) (in which $\varepsilon_{G_{\max,f}} \sim \mathcal{N}(0, \sigma_{G_{\max,f}})$, $\varepsilon_{G_{\max,p}} \sim \mathcal{N}(0, \sigma_{G_{\max,p}})$ and $\varepsilon_{G_{\max,s}} \sim \mathcal{N}(0, \sigma_{G_{\max,s}})$).

As shown in Fig. 1, competitive parameter (α) was modelled using an equation of the form:

$$\alpha_{f,c} = \alpha_{0\text{intra},f} C + \alpha_{0\text{inter},f} (1 - C) - \alpha_t t_f + \alpha_e t_c + \alpha_d |t_c - t_f| \quad (4)$$

in which $\alpha_{0\text{intra},f}$ and $\alpha_{0\text{inter},f}$ are (respectively) intraspecific and average interspecific trait independent competition for the focal species (f), modelled each with a normally distributed random effect of species (f) and normally distributed random effect of data set (s) (such as $\alpha_{0\text{intra},f} = \alpha_{0\text{intra}} + \varepsilon_{\alpha_{0\text{intra},f}} + \varepsilon_{\alpha_{0\text{intra},s}}$, in which $\varepsilon_{\alpha_{0\text{intra},f}} \sim \mathcal{N}(0, \sigma_{\alpha_{0\text{intra},f}})$ and $\varepsilon_{\alpha_{0\text{intra},s}} \sim \mathcal{N}(0, \sigma_{\alpha_{0\text{intra},s}})$), and replacing intra by inter gives the expressions for $\alpha_{0\text{inter},f}$. C is a binary variable taking the value one for $f=c$ (conspecific) and zero for $f \neq c$ (heterospecific), α_t is the tolerance of competition by the focal species, that is, change in competition tolerance due to traits (t_f) of the focal tree with a normally distributed random effect of data set (s) included ($\varepsilon_{\alpha_t,s} \sim \mathcal{N}(0, \sigma_{\alpha_t,s})$). α_e is the competitive effect, that is, change in competition effect due to traits (t_c) of the competitor tree with a normally distributed random effect of data set (s) included ($\varepsilon_{\alpha_e,s} \sim \mathcal{N}(0, \sigma_{\alpha_e,s})$), and α_d is the effect of trait dissimilarity, that is, change in competition due to absolute distance between traits $|t_c - t_f|$ with a normally distributed random effect of data set (s) included ($\varepsilon_{\alpha_d,s} \sim \mathcal{N}(0, \sigma_{\alpha_d,s})$).

Our decomposition of the competition parameter (α) into trait-based processes builds on previous studies. In one of the first studies, Uriarte *et al.*⁸ modelled α as $\alpha = \alpha_0 + \alpha_d |t_f - t_c|$. Then, Kunstler *et al.*⁹ used two different models: $\alpha = \alpha_0 + \alpha_d |t_f - t_c|$ or $\alpha = \alpha_0 + \alpha_h |t_f - t_c|$. Finally, Lasky *et al.*¹¹ developed a single model including multiple processes as $\alpha = \alpha_0 + t_f + \alpha_h (t_f - t_c) + \alpha_d |t_f - t_c|$. Here we extended the approach of this most recent study¹¹ by splitting $\alpha_h(t_f - t_c)$ into $\alpha_t t_f + \alpha_e t_c$ (which is equivalent to the hierarchical distance if $\alpha_t = -\alpha_e$) and including two α_0 , one for intraspecific and one for interspecific competition.

To simplify the estimation, equation (4) was combined with the basal area of each competing species to relate the parameters directly to the community weighted means of the different trait variables as:

$$\sum_{c=1}^{N_i} \alpha_{c,f} B_{i,c,p,s} = \alpha_{0,f,\text{intra}} B_{i,f} + \alpha_{0,f,\text{inter}} B_{i,\text{het}} - \alpha_t t_f B_{i,\text{tot}} + \alpha_e B_{i,t_c} + \alpha_d B_{i,|t_c - t_f|} \quad (5)$$

Where $B_{i,\text{het}} = \sum_{c \neq f} B_{i,c}$ is the sum of basal area of heterospecific competitors (het), $B_{i,\text{tot}} = B_{i,f} + B_{i,\text{het}}$ is the sum of basal area of all competitors, $B_{i,t_c} = \sum_{c=1}^{N_i} t_c \times B_{i,c}$ and $B_{i,|t_c - t_f|} = \sum_{c=1}^{N_i} |t_c - t_f| \times B_{i,c}$. N_i is the number of species in the local neighbourhood of the tree (i) (note that the indices p and s for plot and data set are not shown here for sake of simplicity).

Estimating separate α_0 for intra and interspecific competition allowed us to account for trait-independent differences in interactions with conspecifics and heterospecifics. We also explored a simpler version of the model where trait-independent competitive effects were pooled (that is, there was a single value for α_0), as previous studies have generally not made this distinction, using the following equation:

$$\alpha_{f,c} = \alpha_{0,f} - \alpha_t t_f + \alpha_e t_c + \alpha_d |t_c - t_f| \quad (6)$$

In this alternative model, any differences between intra and interspecific competition do enter into trait dissimilarity effects, with a trait dissimilarity of zero attached to them. This may lead to an overestimation of the trait dissimilarity effect. Results for this model are presented in Extended Data Fig. 4.

Equations (2)–(4) were then fitted to empirical estimates of growth based on change in diameter between census t and $t+1$ (respectively at year y_t and y_{t+1}), given by

$$G_{i,f,p,s,t} = 0.25\pi(D_{i,f,p,s,t+1}^2 - D_{i,f,p,s,t}^2)/(y_{t+1} - y_t) \quad (7)$$

To estimate standardised coefficients (one type of standardised effect size)³¹, response and explanatory variables were standardized (divided by their standard deviations) before analysis. Trait and diameter were also centred to facilitate convergence. The models were fitted using the *lmer* routine in the *lme4* package³² in the R statistical environment³³. We fitted two versions of each model. In the first version parameters m_0 , m_1 , α_0 , α_t , α_e , α_d were estimated as constant across all biomes. In the second version, we allowed different fixed estimates of these parameters for each biome. This enabled us to explore variation among biomes. Because some biomes had few observations, we merged those with biomes with similar climates. Tundra was merged with taiga, tropical rainforest and tropical seasonal forest were merged into tropical forest, and deserts were not included in this final analysis as too few plots were available. To evaluate whether our results were robust to the random effect structure we also explored a model with a random effect attached to parameters both for the data set and for a local ecoregion using the Köppen–Geiger ecoregion³⁴ (see Supplementary Results).

Estimating the effect of traits on the average differences between intra and interspecific competition. Differences between inter and intraspecific competition have long been considered key to community assembly and species coexistence^{12,35–38}. Our estimated growth model allowed us to estimate the average inter and intraspecific competition from trait-independent and trait-dependent processes. For any combination of two trait values t_i and t_j , we can predict the interspecific (α_{t_i,t_j} and α_{t_j,t_i}) and intraspecific (α_{t_i,t_i} and α_{t_j,t_j}) competition parameters for a typical species by leaving out the random species effects in equation (4). We can then estimate the average differences between interspecific and intraspecific competition over all trait values combinations using the following expression:

$$\frac{(\alpha_{t_i,t_j} - \alpha_{t_i,t_i}) + (\alpha_{t_j,t_i} - \alpha_{t_j,t_j})}{2} \quad (8)$$

Substituting in from equation (4) (leaving out the species random effect) this simplifies as:

$$\alpha_{0\text{inter}} - \alpha_{0\text{intra}} + \alpha_d |t_j - t_i| \quad (9)$$

Thus, the average differences between inter and intraspecific competition are affected only by the difference between $\alpha_{0\text{intra}}$ and $\alpha_{0\text{inter}}$ and by trait dissimilarity via α_d (see Extended Data Fig. 2 for the results).

Growth data. Our main objective was to collate data sets spanning the dominant forest biomes of the world. Data sets were included if they allowed both growth of individual trees and the local abundance of competitors to be estimated, and if they had good (>40%) coverage for at least one of the traits of interest (SLA, wood density and maximum height).

The data sets collated fell into two broad categories: (1) national forest inventories (NFI), in which trees above a given diameter were sampled in a network of small plots (often on a regular grid) covering the country (references for NFI data used^{39–48}); (2) large permanent plots (LPP) ranging in size from 0.5 to 50 ha, in which the x – y coordinates of all trees above a given diameter were recorded (references for LPP data used refs 49–56). LPP were mostly located in tropical regions.

The minimum diameter of recorded trees varied among sites from 1 to 12 cm. To allow comparison between data sets, we restricted our analysis to trees greater than 10 cm. Moreover, we excluded from the analysis any plots with harvesting during the growth measurement period, that were identified as plantations, or that overlapped a forest edge. Finally, we randomly selected only two consecutive census dates per plot or quadrat to avoid having to account for repeated measurements (less than one-third of the data had repeated measurements). Because human and natural disturbances are present in all of these forests (see Supplementary Information), they probably all experience successional dynamics (as indicated by the forest age distribution available in some of these sites in Supplementary Information). See Supplementary Information and Extended Data Table 2 for more details on individual data sets.

Basal area growth was estimated from diameter measurements recorded between the two censuses. For the French NFI, these data were obtained from short tree cores. For all other data sets, diameter at breast height (D) of each individual was recorded at multiple census dates. We excluded trees (1) with extreme positive or negative diameter growth measurements, following criteria developed at the BCI site⁵⁰ (see the R package CTFS R), (2) that were palms or tree ferns, or (3) that were measured at different heights in two consecutive censuses.

For each individual tree, we estimated the local abundance of competitor species as the sum of basal area for all individuals > 10 cm diameter within a specified neighbourhood. For LPPs, we defined the neighbourhood as being a circle with a 15 m radius. This value was selected based on previous studies showing the maximum radius of interaction to lie in the range 10–20 m (refs 8, 19). To avoid edge effects, we also excluded trees less than 15 m from the edge of a plot. To account for variation of abiotic conditions within the LPPs, we divided plots into regularly spaced 20×20 m quadrats and included a random quadrat effect in the model (see above).

For NFI data coordinates of individual trees within plots were generally not available, thus neighbourhoods were defined based on plot size. In the NFI from the United States, four sub-plots of 7.35 m located within 20 m of one another were measured. We grouped these sub-plots to give a single estimate of the local competitor abundance. Thus, the neighbourhoods used in the competition analysis ranged in size from 10 to 25 m radius, with most plots from 10 to 15 m radius. We included variation in neighbourhood size in the constant w_j to compute competitor basal area in $\text{m}^2 \text{ha}^{-1}$.

We extracted MAT and MAP from the WorldClim data base⁵⁷ using the plot latitude and longitude (see Extended Data Fig. 1 for plot locations). MAT and MAP data were then used to classify plots into biomes, using the diagram provided by ref. 58 (modified from ref. 59).

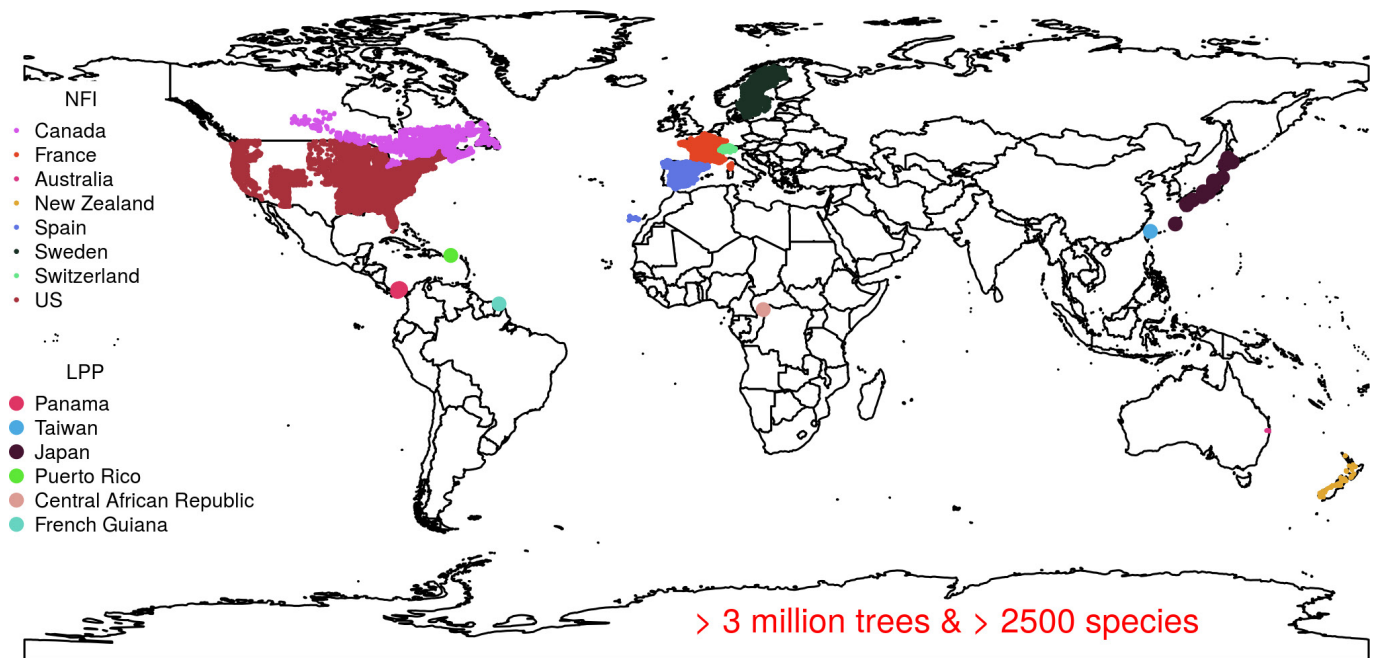
Traits. Data on species functional traits were extracted from existing sources. We focused on wood density, species SLA and maximum height, because these traits have previously been related to competitive interactions and are available for large numbers of species^{8,9,11,15,30} (see Extended Data Tables 3 and 4 for trait coverage and their correlations). Where available, we used data collected locally (references for the local trait data used in this analysis include refs 15, 51, 60–62); otherwise we sourced data from the TRY trait data base¹⁸ (references for the data extracted from the TRY database used in this analysis include refs 2, 3, 15, 63–130). Local data were available for most tropical sites and species (see Supplementary Information). Several of the NFI data sets also provided tree height measurements, from which we computed a species' maximum height as the 99% quantile of observed values (for France, USA, Spain and Switzerland). For Sweden, we used the estimate from the French data set and for Canada we used the estimate from the USA data set. Otherwise, we extracted height measurements from the TRY database. We were not able to account for trait variability within species.

For each focal tree, our approach required us to also account for the traits of all competitors present in the neighbourhood. Most of our plots had good coverage of competitors, but inevitably there were some trees where trait data were lacking. In these cases we estimated trait data as follows. If possible, we used the genus mean, and if no genus data was available, we used the mean of the species present in the country. However, we restricted our analysis to plots where (1) the percentage of basal area contributed by trees with no species level trait data was less than 10%, and (2) the percentage of basal area of trees with neither species nor genus level trait data was less than 5%.

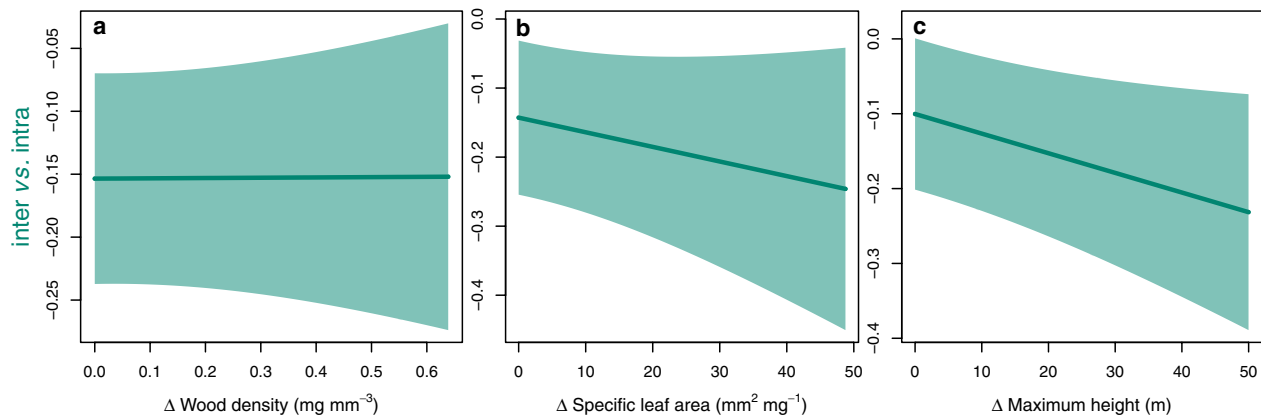
29. Canham, C. D. *et al.* Neighborhood analyses of canopy tree competition along environmental gradients in New England forests. *Ecol. Appl.* **16**, 540–554 (2006).
30. Rüger, N., Wirth, C., Wright, S. J. & Condit, R. Functional traits explain light and size response of growth rates in tropical tree species. *Ecology* **93**, 2626–2636 (2012).
31. Schielzeth, H. Simple means to improve the interpretability of regression coefficients: Interpretation of regression coefficients. *Methods Ecol. Evol.* **1**, 103–113 (2010).

32. Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* **67**, 1–48 (2014).
33. R Development Core Team. *R: A Language and Environment for Statistical Computing* <http://www.R-project.org/> (R Foundation for Statistical Computing, 2014).
34. Kriticos, D. J. *et al.* CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods Ecol. Evol.* **3**, 53–64 (2012).
35. Connell, J. H. On the prevalence and relative importance of interspecific competition: Evidence from field experiments. *Am. Nat.* **122**, 661–696 (1983).
36. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343–366 (2000).
37. Chesson, P. in *Ecological Systems* (ed. Leemans, R.) 223–256 (Springer, 2012).
38. Godoy, O. & Levine, J. M. Phenology effects on invasion success: insights from coupling field experiments to coexistence theory. *Ecology* **95**, 726–736 (2014).
39. IGN Inventaire Forestier; <http://inventaire-forestier.ign.fr/spip/spip.php?rubrique153>.
40. Kooyman, R., Rossetto, M., Allen, C. & Cornwell, W. Australian tropical and subtropical rain forest community assembly: phylogeny, functional biogeography, and environmental gradients. *Biotropica* **44**, 668–679 (2012).
41. New Zealand National Vegetation Survey Databank; <https://nvs.landcareresearch.co.nz/>.
42. Wiser, S. K., Bellingham, P. J. & Burrows, L. E. Managing biodiversity information: development of New Zealand's National Vegetation Survey databank. *N. Z. J. Ecol.* **25**, 1–17 (2001).
43. Ministerio de Agricultura, Alimentación y Medio Ambiente. *Inventario Forestal Nacional*; <http://www.magrama.gob.es/es/desarrollo-rural/temas/politica-forestal/inventario-cartografia/inventario-forestal-nacional/default.aspx>.
44. Villaseca, R. & Diaz, R. *Segundo Inventario Forestal Nacional (1986–1996)* (Ministerio de Medio Ambiente, 1998).
45. Villanueva, J. *Tercer Inventario Forestal Nacional (1997–2007)* (Ministerio de Medio Ambiente, 2004).
46. Fridman, J. & Stahl, G. A three-step approach for modelling tree mortality in Swedish forests. *Scand. J. For. Res.* **16**, 455–466 (2001).
47. Swiss National Forest Inventory (NFI); <http://www.lfi.ch/index-en.php>.
48. USDA Forest Inventory and Analysis National Program; <http://www.fia.fs.fed.us/tools-data/>.
49. Condit, R., Engelbrecht, B. M. J., Pino, D., Perez, R. & Turner, B. L. Species distributions in response to individual soil nutrients and seasonal drought across a community of tropical trees. *Proc. Natl Acad. Sci. USA* **110**, 5064–5068 (2013).
50. Condit, R., Hubbell, S. P. & Foster, R. B. Mortality and growth of a commercial hardwood 'el catio', *Prioria copaifera*, in Panama. *For. Ecol. Manage.* **62**, 107–122 (1993).
51. Lasky, J. R., Sun, I., Su, S.-H., Chen, Z.-S. & Keitt, T. H. Trait-mediated effects of environmental filtering on tree community dynamics. *J. Ecol.* **101**, 722–733 (2013).
52. Ishihara, M. *et al.* Forest stand structure, composition, and dynamics in 34 sites over Japan. *Ecol. Res.* **26**, 1007–1008 (2011).
53. Thompson, J. *et al.* Land use history, environment, and tree composition in a tropical forest. *Ecol. Appl.* **12**, 1344–1363 (2002).
54. Ouedraogo, D.-Y., Mortier, F., Gourlet-Fleury, S., Freycon, V. & Picard, N. Slow-growing species cope best with drought: evidence from long-term measurements in a tropical semi-deciduous moist forest of Central Africa. *J. Ecol.* **101**, 1459–1470 (2013).
55. Hérault, B., Ouallet, J., Blanc, L., Wagner, F. & Baraloto, C. Growth responses of neotropical trees to logging gaps. *J. Appl. Ecol.* **47**, 821–831 (2010).
56. Hérault, B. *et al.* Functional traits shape ontogenetic growth trajectories of rain forest tree species. *J. Ecol.* **99**, 1431–1440 (2011).
57. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
58. Ricklefs, R. E. *The Economy of Nature* (WH Freeman, 2001).
59. Whittaker, R. H. *Communities and Ecosystems* (Macmillan, 1970).
60. Swenson, N. G. *et al.* Temporal turnover in the composition of tropical tree communities: functional determinism and phylogenetic stochasticity. *Ecology* **93**, 490–499 (2012).
61. Gourlet-Fleury, S. *et al.* Environmental filtering of dense-wooded species controls above-ground biomass stored in African moist forests. *J. Ecol.* **99**, 981–990 (2011).
62. Baraloto, C. *et al.* Decoupled leaf and stem economics in rain forest trees. *Ecol. Lett.* **13**, 1338–1347 (2010).
63. Ackerly, D. D. & Cornwell, W. K. A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecol. Lett.* **10**, 135–145 (2007).
64. Castro-Diez, P., Puyravaud, J., Cornelissen, J. & Villar-Salvador, P. Stem anatomy and relative growth rate in seedlings of a wide range of woody plant species and types. *Oecologia* **116**, 57–66 (1998).
65. Cornelissen, J. An experimental comparison of leaf decomposition rates in a wide range of temperate plant species and types. *J. Ecol.* **84**, 573–582 (1996).
66. Cornelissen, J., Diez, P. C. & Hunt, R. Seedling growth, allocation and leaf attributes in a wide range of woody plant species and types. *J. Ecol.* **84**, 755–765 (1996).
67. Cornelissen, J., Werger, M., Castro-Diez, P., Van Rheenen, J. & Rowland, A. Foliar nutrients in relation to growth, allocation and leaf traits in seedlings of a wide range of woody plant species and types. *Oecologia* **111**, 460–469 (1997).
68. Cornelissen, J. *et al.* Leaf digestibility and litter decomposability are related in a wide range of subarctic plant species and types. *Funct. Ecol.* **18**, 779–786 (2004).

69. Cornelissen, J. *et al.* Functional traits of woody plants: correspondence of species rankings between field adults and laboratory-grown seedlings? *J. Veg. Sci.* **14**, 311–322 (2003).
70. Cornwell, W. K. & Ackerly, D. D. Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecol. Monogr.* **79**, 109–126 (2009).
71. Cornwell, W. K., Schwillk, L. D. W. & Ackerly, D. D. A trait-based test for habitat filtering: convex hull volume. *Ecology* **87**, 1465–1471 (2006).
72. Cornwell, W., Bhaskar, R., Sack, L. & Cordell, S. Adjustment of structure and function of Hawaiian *Metrosideros polymorpha* at high vs. low precipitation. *Funct. Ecol.* **21**, 1063–1071 (2007).
73. Cornwell, W. K. *et al.* Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecol. Lett.* **11**, 1065–1071 (2008).
74. Diaz, S. *et al.* The plant traits that drive ecosystems: evidence from three continents. *J. Veg. Sci.* **15**, 295–304 (2004).
75. Fonseca, C. R., Overton, J. M., Collins, B. & Westoby, M. Shifts in trait-combinations along rainfall and phosphorus gradients. *J. Ecol.* **88**, 964–977 (2000).
76. Fortunel, C. *et al.* Leaf traits capture the effects of land use changes and climate on litter decomposability of grasslands across Europe. *Ecology* **90**, 598–611 (2009).
77. Freschet, G. T., Cornelissen, J. H., Van Logtestijn, R. S. & Aerts, R. Evidence of the 'plant economics spectrum' in a subarctic flora. *J. Ecol.* **98**, 362–373 (2010).
78. Freschet, G. T., Cornelissen, J. H., van Logtestijn, R. S. & Aerts, R. Substantial nutrient resorption from leaves, stems and roots in a subarctic flora: what is the link with other resource economics traits? *New Phytol.* **186**, 879–889 (2010).
79. Garnier, E. *et al.* Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: a standardized methodology and lessons from an application to 11 European sites. *Ann. Bot. (Lond.)* **99**, 967–985 (2007).
80. Green, W. USDA PLANTS compilation, version 1, 09-02-02; <http://bricol.net/downloads/data/PLANTSdatabase/> (2009).
81. Han, W., Fang, J., Guo, D. & Zhang, Y. Leaf nitrogen and phosphorus stoichiometry across 753 terrestrial plant species in China. *New Phytol.* **168**, 377–385 (2005).
82. He, J.-S. *et al.* A test of the generality of leaf trait relationships on the Tibetan Plateau. *New Phytol.* **170**, 835–848 (2006).
83. He, J.-S. *et al.* Leaf nitrogen: phosphorus stoichiometry across Chinese grassland biomes. *Oecologia* **155**, 301–310 (2008).
84. Hoof, J., Sack, L., Webb, D. T. & Nilsen, E. T. Contrasting structure and function of pubescent and glabrous varieties of Hawaiian *Metrosideros polymorpha* (Myrtaceae) at high elevation. *Biotropica* **40**, 113–118 (2008).
85. Kattge, J., Knorr, W., Raddatz, T. & Wirth, C. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Glob. Change Biol.* **15**, 976–991 (2009).
86. Kleyer, M. *et al.* The LEDA traitbase: a database of life-history traits of the Northwest European flora. *J. Ecol.* **96**, 1266–1274 (2008).
87. Kurokawa, H. & Nakashizuka, T. Leaf herbivory and decomposability in a Malaysian tropical rain forest. *Ecology* **89**, 2645–2656 (2008).
88. Laughlin, D. C., Leppert, J. J., Moore, M. M. & Sieg, C. H. A multi-trait test of the leaf-height-seed plant strategy scheme with 133 species from a pine forest flora. *Funct. Ecol.* **24**, 493–501 (2010).
89. Martin, R. E., Asner, G. P. & Sack, L. Genetic variation in leaf pigment, optical and photosynthetic function among diverse phenotypes of *Metrosideros polymorpha* grown in a common garden. *Oecologia* **151**, 387–400 (2007).
90. McDonald, P., Fonseca, C., Overton, J. & Westoby, M. Leaf-size divergence along rainfall and soil-nutrient gradients: is the method of size reduction common among clades? *Funct. Ecol.* **17**, 50–57 (2003).
91. Medlyn, B. E. *et al.* Effects of elevated [CO₂] on photosynthesis in European forest species: a meta-analysis of model parameters. *Plant Cell Environ.* **22**, 1475–1495 (1999).
92. Medlyn, B. E. & Jarvis, P. G. Design and use of a database of model parameters from elevated [CO₂] experiments. *Ecol. Modell.* **124**, 69–83 (1999).
93. Medlyn, B. *et al.* Stomatal conductance of forest species after long-term exposure to elevated CO₂ concentration: a synthesis. *New Phytol.* **149**, 247–264 (2001).
94. Messier, J., McGill, B. J. & Lechowicz, M. J. How do traits vary across ecological scales? A case for trait-based ecology. *Ecol. Lett.* **13**, 838–848 (2010).
95. Moles, A. T. *et al.* Factors that shape seed mass evolution. *Proc. Natl Acad. Sci. USA* **102**, 10540–10544 (2005).
96. Moles, A. T. *et al.* A brief history of seed size. *Science* **307**, 576–580 (2005).
97. Moles, A. T., Falster, D. S., Leishman, M. R. & Westoby, M. Small-seeded species produce more seeds per square metre of canopy per year, but not per individual per lifetime. *J. Ecol.* **92**, 384–396 (2004).
98. Niinemets, Ü. Global-scale climatic controls of leaf dry mass per area, density, and thickness in trees and shrubs. *Ecology* **82**, 453–469 (2001).
99. Niinemets, Ü. Research review. components of leaf dry mass per area—thickness and density—alter leaf photosynthetic capacity in reverse directions in woody plants. *New Phytol.* **144**, 35–47 (1999).
100. Ogaya, R. & Peñuelas, J. Experimental drought in a holm oak forest: different photosynthetic response of the two dominant species, *Quercus ilex* and *Phillyrea latifolia*. *Environ. Exp. Bot.* **50**, 137–148 (2003).
101. Ogaya, R. & Peñuelas, J. Contrasting foliar responses to drought in *Quercus ilex* and *Phillyrea latifolia*. *Biol. Plant.* **50**, 373–382 (2006).
102. Ogaya, R. & Peñuelas, J. Tree growth, mortality, and above-ground biomass accumulation in a holm oak forest under a five-year experimental field drought. *Plant Ecol.* **189**, 291–299 (2007).
103. Ogaya, R. & Peñuelas, J. Tree growth, mortality, and above-ground biomass accumulation in a holm oak forest under a five-year experimental field drought. *Plant Ecol.* **189**, 291–299 (2007).
104. Onoda, Y. *et al.* Global patterns of leaf mechanical properties. *Ecol. Lett.* **14**, 301–312 (2011).
105. Ordoñez, J. C. *et al.* Plant strategies in relation to resource supply in mesic to wet environments: does theory mirror nature? *Am. Nat.* **175**, 225–239 (2010).
106. Ordoñez, J. C. *et al.* Leaf habit and woodiness regulate different leaf economy traits at a given nutrient supply. *Ecology* **91**, 3218–3228 (2010).
107. Pakeman, R. J. *et al.* Impact of abundance weighting on the response of seed traits to climate and land use. *J. Ecol.* **96**, 355–366 (2008).
108. Pakeman, R. J., Lepš, J., Kleyer, M., Lavorel, S. & Garnier, E. Relative climatic, edaphic and management controls of plant functional trait signatures. *J. Veg. Sci.* **20**, 148–159 (2009).
109. Peñuelas, J. *et al.* Faster returns on 'leaf economics' and different biogeochemical niche in invasive compared with native plant species. *Glob. Change Biol.* **16**, 2171–2185 (2010).
110. Peñuelas, J. *et al.* Higher allocation to low cost chemical defenses in invasive species of Hawaii. *J. Chem. Ecol.* **36**, 1255–1270 (2010).
111. Poorter, L. & Bongers, F. Leaf traits are good predictors of plant performance across 53 rain forest species. *Ecology* **87**, 1733–1743 (2006).
112. Poorter, L. Leaf traits show different relationships with shade tolerance in moist versus dry tropical forests. *New Phytol.* **181**, 890–900 (2009).
113. Poorter, H., Niinemets, Ü., Poorter, L., Wright, I. J. & Villar, R. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytol.* **182**, 565–588 (2009).
114. Preston, K. A., Cornwell, W. K. & DeNoyer, J. L. Wood density and vessel traits as distinct correlates of ecological strategy in 51 California coast range angiosperms. *New Phytol.* **170**, 807–818 (2006).
115. Pyankov, V. I., Kondratyuk, A. V. & Shipley, B. Leaf structure and specific leaf mass: the alpine desert plants of the Eastern Pamirs, Tajikistan. *New Phytol.* **143**, 131–142 (1999).
116. Quested, H. M. *et al.* Decomposition of sub-arctic plants with differing nitrogen economies: a functional role for hemiparasites. *Ecology* **84**, 3209–3221 (2003).
117. Reich, P. B. *et al.* Scaling of respiration to nitrogen in leaves, stems and roots of higher land plants. *Ecol. Lett.* **11**, 793–801 (2008).
118. Reich, P. B., Oleksyn, J. & Wright, I. J. Leaf phosphorus influences the photosynthesis–nitrogen relation: a cross-biome analysis of 314 species. *Oecologia* **160**, 207–212 (2009).
119. Sack, L. Responses of temperate woody seedlings to shade and drought: do trade-offs limit potential niche differentiation? *Oikos* **107**, 110–127 (2004).
120. Sack, L., Tyree, M. T. & Holbrook, N. M. Leaf hydraulic architecture correlates with regeneration irradiance in tropical rainforest trees. *New Phytol.* **167**, 403–413 (2005).
121. Sack, L., Melcher, P. J., Liu, W. H., Middleton, E. & Pardee, T. How strong is intracranopy leaf plasticity in temperate deciduous trees? *Am. J. Bot.* **93**, 829–839 (2006).
122. Sardans, J., Peñuelas, J. & Ogaya, R. Drought-induced changes in C and N stoichiometry in a *Quercus ilex* Mediterranean forest. *For. Sci.* **54**, 513–522 (2008).
123. Sardans, J., Peñuelas, J., Prieto, P. & Estiarte, M. Changes in Ca, Fe, Mg, Mo, Na, and S content in a Mediterranean shrubland under warming and drought. *J. Geophys. Res.* **113**, G03039 (2008).
124. Shipley, B. & Vu, T.-T. Dry matter content as a measure of dry matter concentration in plants and their parts. *New Phytol.* **153**, 359–364 (2002).
125. Soudzilovskaia, N. A. *et al.* Functional traits predict relationship between plant abundance dynamic and long-term climate warming. *Proc. Natl Acad. Sci. USA* **110**, 18180–18184 (2013).
126. Willis, C. G. *et al.* Phylogenetic community structure in Minnesota oak savanna is influenced by spatial extent and environmental variation. *Ecography* **33**, 565–577 (2010).
127. Wilson, K. B., Baldocchi, D. D. & Hanson, P. J. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. *Tree Physiol.* **20**, 565–578 (2000).
128. Wright, I. J. *et al.* Relationships among ecologically important dimensions of plant trait variation in seven neotropical forests. *Ann. Bot. (Lond.)* **99**, 1003–1015 (2007).
129. Wright, I. J. *et al.* Irradiance, temperature and rainfall influence leaf dark respiration in woody plants: evidence from comparisons across 20 sites. *New Phytol.* **169**, 309–319 (2006).
130. Zanne, A. E. *et al.* Angiosperm wood structure: global patterns in vessel anatomy and their relation to wood density and potential conductivity. *Am. J. Bot.* **97**, 207–215 (2010).
131. South, A. rworldmap, a new R package for mapping global data. *Rem. J.* **3**, 35–43 (2011).
132. Nakagawa, S. & Hanson, P. J. A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**, 133–142 (2013).
133. Johnson, P. C. D. Extension of Nakagawa and Schielzeth's R²GLMM to random slopes models. *Methods Ecol. Evol.* **5**, 944–946 (2014).
134. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach* (Springer-Verlag, 2002).

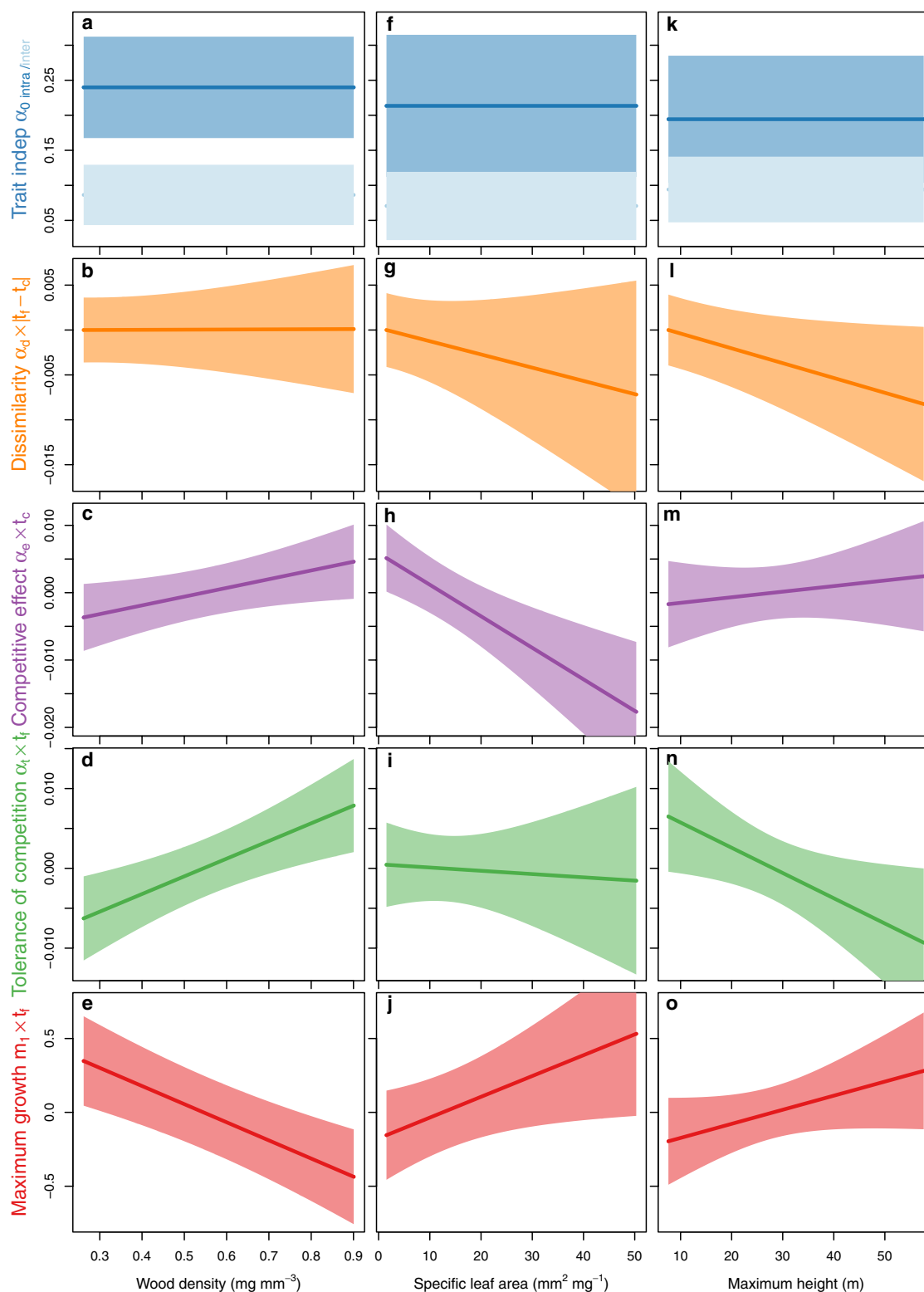


Extended Data Figure 1 | Map of the plot locations of all data sets analysed. LPP plots are represented with a large points and NFI plots with small points (the Panama data set comprises both a 50 ha plot and a network of 1 ha plots). The world map is from the R package *rworldmap*¹³¹ using Natural Earth data.



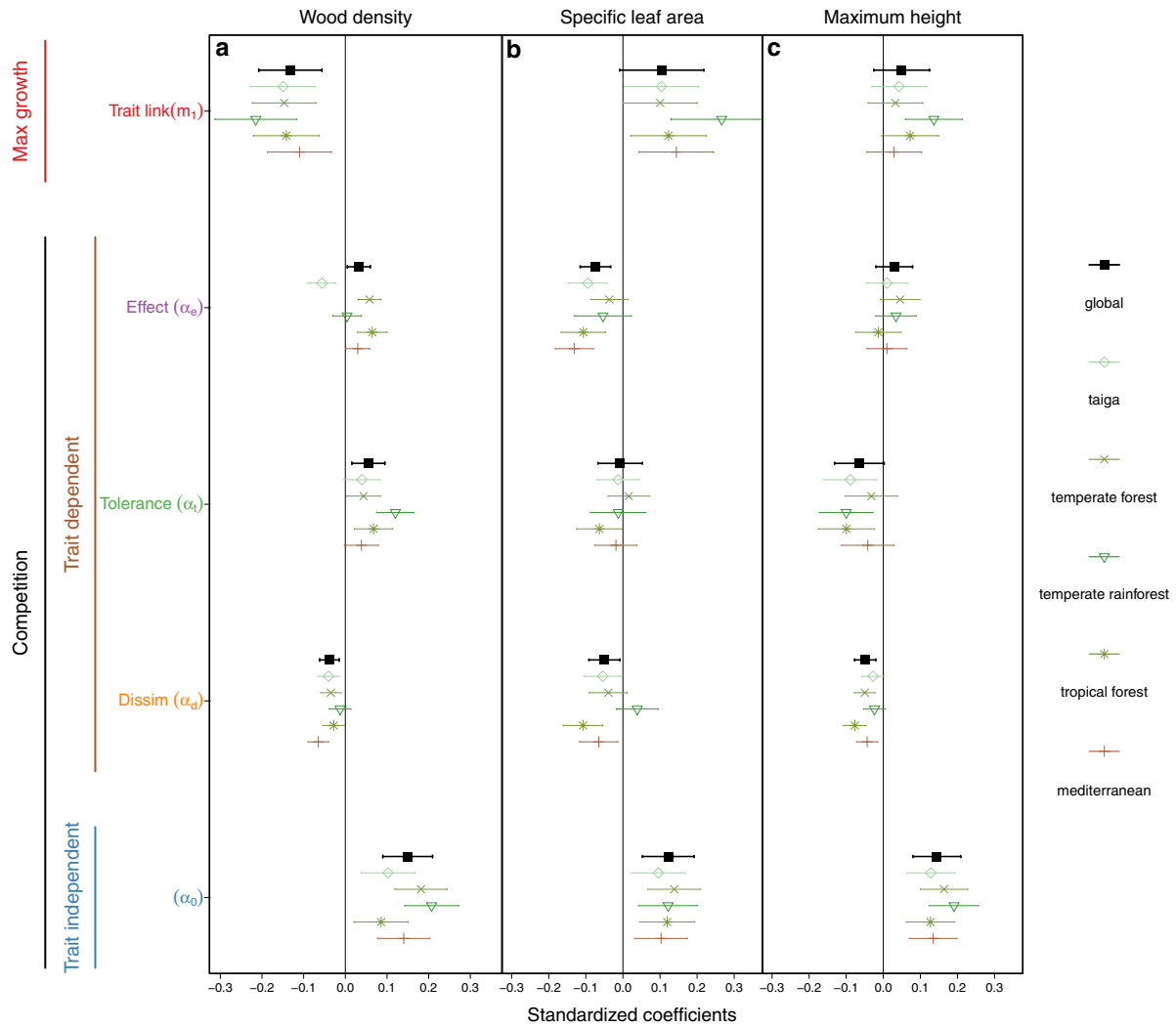
Extended Data Figure 2 | Average difference between interspecific and intraspecific competition predicted with estimates of trait-independent and trait-dependent processes influencing competition for models fitted for each trait. a–c, Models were fitted for wood density (a), SLA

(b) or maximum height (c). The average differences between interspecific and intraspecific competition are influenced by $\alpha_{0\text{intra}}$, $\alpha_{0\text{inter}}$ and α_d coefficients (see Methods for details). Negative values indicate that intraspecific competition is stronger than interspecific competition.



Extended Data Figure 3 | Variation of trait-independent inter and intraspecific competition, trait dissimilarity ($|t_f - t_c| \times \alpha_d$), competitive effect ($t_c \times \alpha_c$), tolerance to competition ($t_f \times \alpha_t$) and maximum growth ($t_f \times m_1$) with wood density, SLA and maximum height. a–o, Wood density (a–e), SLA (f–j) and maximum height (k–o).

Trait varied from their quantile at 5% to their quantile at 95%. The shaded area represents the 95% confidence interval of the prediction (including uncertainty associated with α_0 or m_0). $\alpha_{0\text{intra}}$ and $\alpha_{0\text{inter}}$ which do not vary with traits, are represented with their associated confidence intervals.



Extended Data Figure 4 | Trait-dependent and trait-independent effects on maximum growth and competition across the globe and their variation among biomes for models without separation of α_0 between intra and interspecific competition for wood density, SLA

and maximum height. a, Wood density. b, SLA. c, Maximum height. See Fig. 2 in the main text for parameters description, and see Fig. 1a in the main text for biome definition.

Extended Data Table 1 | Standardized coefficient estimates from models fitted for each trait

	Wood density	SLA	Maximum height
m_0	0.016 (0.127)	-0.087 (0.132)	0.084 (0.089)
γ	0.418 (0.011)	0.401 (0.012)	0.42 (0.01)
m_1	-0.149 (0.036)	0.119 (0.057)	0.063 (0.04)
m_2	0.111 (0.003)	0.093 (0.003)	0.081 (0.002)
m_3	0.053 (0.002)	0.056 (0.003)	0.048 (0.002)
$\alpha_{0 \text{ intra}}$	0.24 (0.037)	0.213 (0.052)	0.194 (0.046)
$\alpha_{0 \text{ inter}}$	0.086 (0.022)	0.071 (0.025)	0.094 (0.024)
α_e	0.034 (0.016)	-0.083 (0.023)	0.017 (0.026)
α_t	0.069 (0.021)	-0.009 (0.033)	-0.071 (0.032)
α_d	0 (0.009)	-0.018 (0.015)	-0.017 (0.008)
R_m^2 *	0.1393	0.1637	0.1429
R_c^2 *	0.7297	0.7593	0.7166
ΔAIC	0	0	0
$\Delta \text{AIC no trait}$	2469	1651	2748

Estimates and their standard error (in brackets) estimated for each trait, R^2 * of models and ΔAIC of the model and of a model with no trait effect. See Methods for explanation of parameters

*We report the conditional and marginal R^2 of the models using the methods of ref. 132, modified by ref. 133. ΔAIC is the difference in the Akaike's information criterion (AIC; as defined by ref. 134) between the model and the best model (lowest AIC). The best-fitting model was identified as the one with a ΔAIC of 0. ΔAIC greater than 10 shows strong support for the best model¹³⁴.

Extended Data Table 2 | Trees data description

set	# of trees	# of species	# of plots/quadrats	% of angiosperm	% of evergreen
Sweden	202480	26	22552	27.0	73.0
New Zealand	53775	117	1415	94.0	99.1
US	1370541	492	59840	63.3	37.2
Canada	495008	75	14983	34.4	64.8
Australia	906	101	63	99.9	92.4
France	184316	127	17611	74.1	28.5
Switzerland	28286	60	2597	36.4	55.2
Spain	418805	122	36462	34.7	81.6
Panama	27089	237	2033	99.8	77.7
French Guiana	46360	712	2157	100.0	83.5
Japan	4658	139	318	72.8	70.0
Taiwan	14701	72	623	92.0	75.3
Puerto Rico	14011	82	399	100.0	99.0
Central African Republic	17638	204	989	99.5	72.4

For each site, the number of individual trees, species and plots in NFI data and quadrats in LPP data, and the percentage of angiosperm and evergreen species are shown.

Extended Data Table 3 | Traits data description

set	% cover SLA	% cover Wood density	% cover Max height
Sweden	99.7	99.6	97.9
New Zealand	99.8	99.5	99.9
US	91.3	94.4	100.0
Canada	99.4	99.4	99.9
Australia	0.0	99.2	100.0
France	99.2	98.9	99.9
Switzerland	96.7	95.1	99.7
Spain	97.3	98.9	100.0
Panama	93.0	93.1	95.4
French Guiana	73.3	73.5	63.5
Japan	99.7	99.7	100.0
Taiwan	99.9	99.3	95.8
Puerto Rico	99.3	99.3	99.3
Central African Republic	40.3	47.1	0.0

The coverage in each site is given with the percentage of species with species level trait data.

Extended Data Table 4 | Species traits pairwise correlations

	Wood density	SLA	Max height
Wood density	1	0.18	-0.04
SLA		1	0.24
Max height			1.00

Pearson's *R* correlations for the three traits.

Earliest hominin occupation of Sulawesi, Indonesia

Gerrit D. van den Bergh^{1,2}, Bo Li¹, Adam Brumm^{3,4}, Rainer Grün³, Dida Yurnaldi^{1,5}, Mark W. Moore⁶, Iwan Kurniawan⁵, Ruly Setiawan^{1,5}, Fachroel Aziz⁵, Richard G. Roberts¹, Suyono⁵, Michael Storey⁷, Erick Setiabudi⁵ & Michael J. Morwood^{1‡}

Sulawesi is the largest and oldest island within Wallacea, a vast zone of oceanic islands separating continental Asia from the Pleistocene landmass of Australia and Papua (Sahul). By one million years ago an unknown hominin lineage had colonized Flores immediately to the south¹, and by about 50 thousand years ago, modern humans (*Homo sapiens*) had crossed to Sahul^{2,3}. On the basis of position, oceanic currents and biogeographical context, Sulawesi probably played a pivotal part in these dispersals⁴. Uranium-series dating of speleothem deposits associated with rock art in the limestone karst region of Maros in southwest Sulawesi has revealed that humans were living on the island at least 40 thousand years ago (ref. 5). Here we report new excavations at Talepu in the Walanae Basin northeast of Maros, where *in situ* stone artefacts associated with fossil remains of megafauna (*Bubalus* sp., *Stegodon* and *Celebochoerus*) have been recovered from stratified deposits that accumulated from before 200 thousand years ago until about 100 thousand years ago. Our findings suggest that Sulawesi, like Flores, was host to a long-established population of archaic hominins, the ancestral origins and taxonomic status of which remain elusive.

In the late 1940s the discovery of 'Palaeolithic' stone artefacts in association with Pleistocene fossil fauna in the Walanae Basin of south Sulawesi⁶ (Fig. 1) led to considerable speculation about the time depth of human occupation of the island^{7,8}. The lithic assemblages comprised cores, choppers and flakes (the 'Cabenge Industry'), and derived from undated surface collections along the eastern side of the Walanae River^{6–8}, which follows the Walanae Depression, an elongated north-south-trending fault-bounded basin (Extended Data Fig. 1). Fossils of several now-extinct species, including two pygmy proboscideans, a giant tortoise and a large endemic suid, *Celebochoerus*, were recovered from the same unstratified contexts^{9,10} and excavations at various sites¹¹. Despite protracted investigations, the stratigraphic context and time range of the 'Cabenge Industry' remained unresolved because of a lack of *in situ* stone artefacts¹¹.

To clarify these issues we conducted surveys in the Cabenge area between 2007 and 2012, leading to the discovery of four new sites with *in situ* stone artefacts in their stratigraphic context. At Talepu, one of the newly discovered sites, we undertook deep-trench excavations. The site is 3 km southeast of Cabenge and 13 km downstream from where the Walanae River leaves its confining valley and enters a widening, actively subsiding floodplain towards the north (Extended Data Fig. 1). During the Pleistocene, east-west compression and wrench faulting along the Walanae fault zone resulted in uplift of the Sengkang anticline and the southern part of the Walanae Depression^{12,13}. In these uplifted areas the folded Pliocene–Pleistocene sedimentary sequences of the Walanae Formation are now exposed¹¹. In the northern part of the Walanae Depression, compressional down-folding facilitated accumulation of fluvio-lacustrine sediments from the Pleistocene to recent times. The Talepu site (4° 22' 06.5" S, 119° 59' 01.7" E) is near the hinge line

between the uplifted southern part and the subsiding northern part of the Walanae Depression.

Our excavations focused on the northernmost hill of an elongated ridge near Talepu village, ~600 m west of the Walanae River (Extended Data Fig. 1d). The Talepu Hill summit lies 32 m above sea level and 18 m above the adjacent Walanae River floodplain. Deposits exposed along this ridge comprise a coarsening-upward sequence of sub-horizontal fluvio-estuarine sand and silt layers overlain by alluvial cobble gravels (Extended Data Fig. 2). Two deep excavations were undertaken at Talepu (trenches T2 and T4) to provide a combined 18.7 m long stratigraphic section exposing five main sedimentary units: units A–E in descending order of depth (Fig. 2).

These excavations revealed the first evidence of *in situ* stone artefacts in securely stratified and dated contexts within the Walanae Basin. The T2 excavation yielded 270 stone artefacts between the surface and 4.2 m depth (Fig. 3a–i, Extended Data Fig. 3b and Supplementary Table 1) which are associated with unit A's high-energy fluvial gravel deposits. Hence most are water-rolled to various degrees, although 21% are in relatively fresh condition. The main source of raw material is coarse- to medium-grained silicified limestone cobbles measuring up to 130 mm in diameter. Most are medium- to large-sized flakes (Supplementary Table 2), with cores comprising 13% of the assemblage. Cores were reduced by hard-hammer blows to one face (42%) or bifacially (58%) from unprepared striking platforms. Core reduction was not intensive although seven cobbles were rotated and subsequent reduction created multiplatform cores. Flakes struck from the cobbles were themselves reduced to one face (60%) or bifacially (40%). Although there is patterning in the flaking techniques, there is little evidence that the stoneworkers were creating tools of specific form; rather, stone-flaking produced sharp-edged flakes for use or as a source for additional flakes.

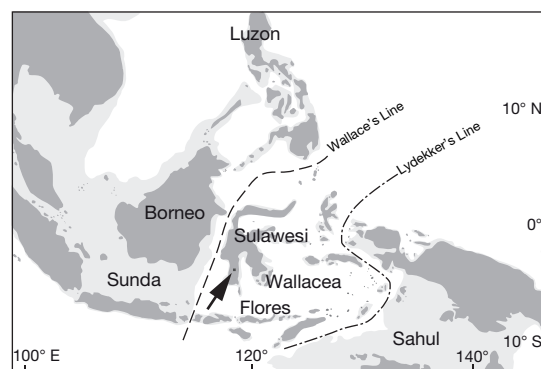


Figure 1 | Sunda, Sahul and Wallacea. Wallacea has two major biogeographical boundaries: Wallace's Line to the west and Lydekker's Line to the east. Exposed land during periods of low sea level (–120 m) is lightly shaded. Talepu Area indicated by an arrow.

¹Centre for Archaeological Science, School of Earth & Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia. ²Naturalis Biodiversity Center, 2333 CR Leiden, The Netherlands. ³Research Centre for Human Evolution, Environmental Futures Research Institute, Griffith University, Nathan, Queensland 4111, Australia. ⁴School of Earth and Environmental Sciences, University of Wollongong, Wollongong, New South Wales 2522, Australia. ⁵Geology Museum Bandung, Geological Agency, Jalan Diponegoro 57, Bandung 40122, Indonesia. ⁶Archaeology, School of Humanities, University of New England, Armidale, New South Wales 2350, Australia. ⁷Quadlab, Natural History Museum of Denmark, University of Copenhagen, Øster Voldgade 5–7, 13 DK-1350 Copenhagen, Denmark. ⁸Deceased.

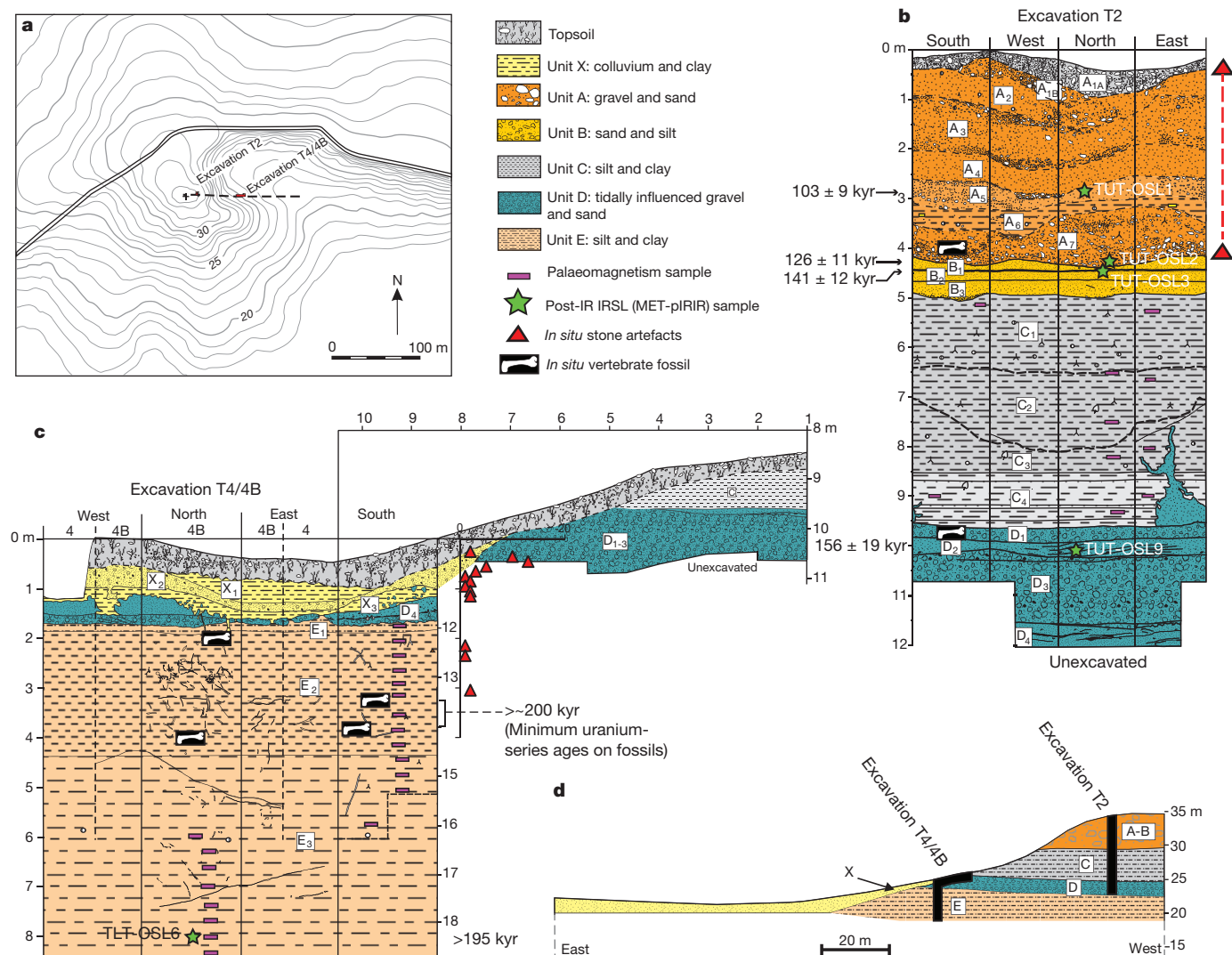


Figure 2 | Talepu excavations T2 and T4. **a**, Site map showing positions of T2 and T4. Dotted line indicates profile shown in **d**; stratigraphy, fossil and stone artefact occurrences, dating sampling horizons and ages for excavations T2 (**b**) and T4 (**c**). Layer key: unit A, conglomerate interval with seven distinct layers: A₁, black topsoil; A₂, sandy pebbly clay; A₃, pebbly sand; A₄ and A₅, coarse pebbly sand; A₆, silty clay; A₇, gravel; unit B, sandy unit: B₁, well-sorted, laminated silty sand; B₂, medium- to fine-grained sand; B₃, well-sorted sand with pebbles near erosive base; unit C, fine-grained sediments: C₁, mottled silty clay; C₂, silty clay;

C₃, mottled silty clay, plant remains; C₄, laminated silty clay with sandy streaks (with sandy intrusion in east baulk); unit D, coarse-grained interval: D₁, well-sorted, mottled pebbly sand; D₂, very coarse cross-bedded sand with rip clasts and mud drapes; D₃, clast-supported gravel with rip clasts; D₄, medium- to fine-grained cross-bedded sand alternating with mud drapes; unit E, fine-grained sediments: E₁, clayey silt; E₂, mottled silts; E₃, clayey silts. Colluvial unit X unconformably overlies unit D. Vertical distribution of *in situ* artefacts in T4 plotted as number of artefacts per 10 cm interval.

From T4 we recovered 41 stone artefacts from the topsoil and colluvium down to a depth of 120 cm. However, four *in situ* silicified limestone artefacts were in exposed older strata within the silt of sub-unit E₂ (Fig. 2c), and provide the stratigraphically earliest evidence for human activity at Talepu. Two are unmodified flakes (2.2–2.4 m depth) (Fig. 3l–m) and two are angular scatter fragments (3.0–3.1 m depth) (Fig. 3j–k), probably created in percussion flaking. The latter are made of a distinctive mottled silicified limestone and appear to have been removed from the same core. The artefacts bear no evidence of water transport; indeed, unit E did not yield any clasts indicative of high-energy water flow.

Only one identifiable fossil was found in T2: a bovid lower molar fragment (4 m depth) (Fig. 3t) that falls just above the size range of the extant lowland anoa, *Bubalus depressicornis* (Extended Data Fig. 3e). T4 yielded eight *Celebochoerus* dental elements (for example, a lower canine; Fig. 3o) and three unidentifiable bone fragments, from the silty interval of sub-unit E₂ (between 3.1–4.0 m below the surface) and just

beneath the lowest stone artefacts. At least some of these fossil remains can be ascribed to one individual (Fig. 3q). A *Stegodon* milk molar fragment (1.9–2.0 m depth) (Fig. 3r) and a dermal scute of a crocodile (3.9–4.0 m) were also recovered.

To constrain the age of the Talepu deposits, we obtained uranium-series ages for teeth and bones excavated from sub-unit E₂ using laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) methods¹⁴ (see Methods: uranium-series dating). Sequential laser spot analyses were undertaken on cross-sections of eight *Celebochoerus* fossils found between 0.2 and 0.5 m below the deepest stone artefacts in the same silty unit (Extended Data Fig. 4a–i). Data sets were combined for each sample and a single age estimate was calculated using the diffusion-absorption-decay model¹⁵. Most of the age results have infinite positive error bounds, so it was only possible to calculate minimum ages¹⁴. The combined uranium-series results (Supplementary Table 3) indicate that the fossil samples are older than ~200 thousand years (kyr). Palaeomagnetic samples from silty layers of units A, C and E have normal magnetic polarities at all sampled levels (Extended Data Fig. 4j

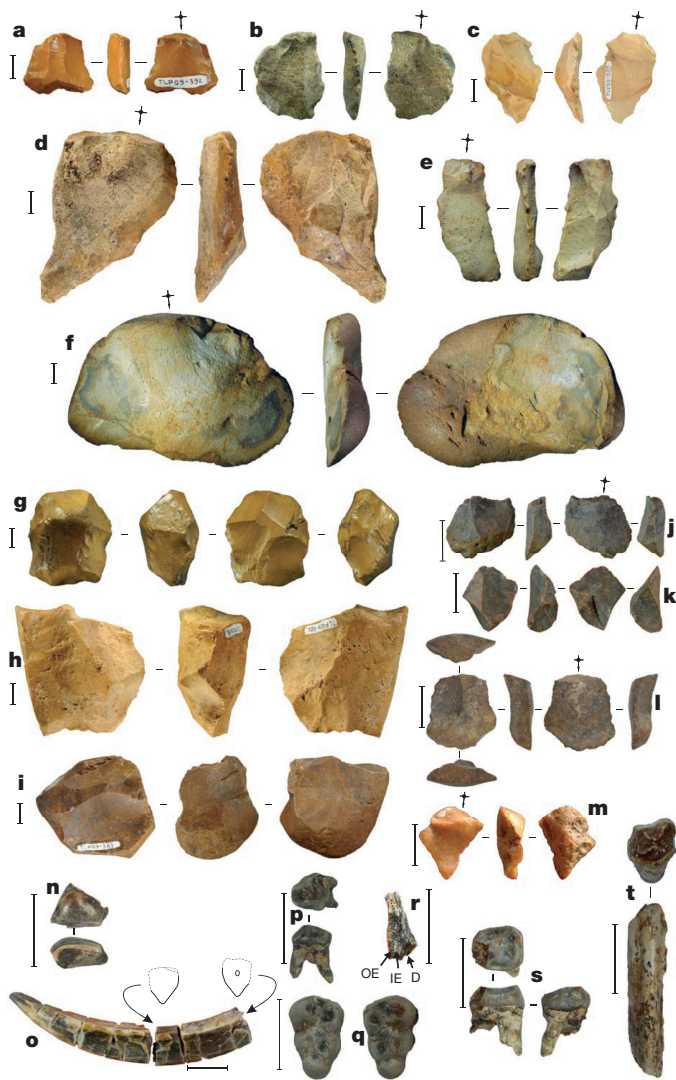


Figure 3 | Finds from Talepu excavations 2 (T2) and 4/4B (T4/T4B). All artefacts shown are silicified limestone. Flakes: **a**, T2-A₇, 3.7–3.8 m; **b**, T2-A₁, 0.4–0.5 m; **c**, T2-A₇, 3.6–3.7 m; **d**, T2-A₂, 0.9 m; **e**, T2-A₁, 0.2–0.3 m; **f**, T2-A₂, 0.8–0.9 m; **j**, T4B-E₂, 3.0–3.1 m; **k**, same material as **j**, but does not refit, T4B-E₂, 3.0–3.1 m; **l**, T4-E₂, 2.38 m; **m**, T4-E₂, 2.2–2.4 m. Cores: **g**, radial core (T2-A₁, 0.1–0.3 m); **h**, core (T2-A₇, 3.6–3.7 m); **i**, core (T2-A₇, 3.6–3.7 m). Fossils ('TLP' numbers refer to the Bandung Geology Museum fossil collection numbers from Talepu): **n**, upper left incisor, *Celebochoerus* sp. (TLP10-F8*, T4-E₂, 3.2–3.4 m); **o**, lower left canine, *Celebochoerus* sp. (TLP10-F1*, T4-E₂, 3.3 m); **p**, upper right third premolar, *Celebochoerus* sp. (TLP10-F5, T4-E₂, 3.3–3.4 m); **q**, upper right (TLP10-F3*) and upper left third molar (TLP10-F4*), *Celebochoerus* sp. (T4-E₂, 3.3–3.4 m); **r**, *Stegodon* molar ridge fragment (D, dentine; OE, outer enamel; IE, inner enamel; OE and IE equal thickness (TLP12-F3, T4B-E₁, 1.9–2 m); **s**, upper left fourth premolar, *Celebochoerus* sp. (TLP10-F2*, T4-E₂, 3.2–3.3 m); **t**, lower left third molar fragment, Bovidae, cf. *Bubalus* sp. (TLP09-F2, T2-A₇, 4–4.1 m). *Fossils used for uranium-series dating. Scale bars: **a**–**i**, 10 mm; **j**–**t**, 20 mm.

and 5). Taken together with the uranium-series results, the fossils are therefore >200 kyr and <780 kyr in age.

To constrain the age of the artefacts, a multiple-elevated-temperature post-infrared infrared stimulated luminescence (MET-pIRIR) dating procedure^{16,17} was applied to potassium-rich feldspar grains extracted from five sediment samples spanning the entire sequence. The four samples analysed from T2 have ages in stratigraphic order 103 ± 9 kyr at 3 m depth to 156 ± 19 kyr at 10 m depth (Fig. 2, Methods: Optical Dating, and Supplementary Table 4). These results suggest that the Talepu cultural sequence ends at ~100 kyr, or possibly earlier

(see Methods). The sediments dated to 156 ± 19 kyr were deposited near the top of unit D, which overlies the sedimentary layer (unit E) from which the deepest artefacts were excavated (more than 3 m below). The oldest securely dated evidence for stone artefacts at Talepu is, therefore, 118 kyr to 194 kyr in age at the 95% confidence interval (2σ), although human occupation of the site clearly occurred earlier given the recovery of artefacts from greater stratigraphic depths (Fig. 2c). Lastly, a sample taken at a depth of 8 m in the lower trench (T4) yielded a minimum age of ~195 kyr. This age estimate is stratigraphically consistent with the MET-pIRIR ages for T2 and with the minimum uranium-series ages of ~200 kyr for the T4 fossil remains from sub-unit E₂.

From our Talepu excavation results it is now possible to conclude that the initial peopling of Sulawesi took place at least 118 thousand years ago (ka). The identity of these early inhabitants is of considerable interest given previous assumptions that Sulawesi was only ever colonized by *H. sapiens*, and currently thought to have arrived in the region by ~50 ka (refs 2, 3, 18–20). The earliest *H. sapiens* skeletal remains from island Southeast Asia are ~45 kyr old^{21,22}; however, modern human fossils dating to ~120 ka occur in the Levant²³, and possibly at a similar time in Southeast Asia²⁴. Although controversial, it is conceivable that *H. sapiens* dispersed soon after their emergence in Africa, spread to the easternmost tip of continental Asia (Sunda) and crossed to Wallacea by ~120 ka. However, early hominins had already reached the more remote and far smaller island of Flores by 1 million years ago (ref. 1), perhaps by accidental drifting on tsunami debris⁴. It is therefore also conceivable that the first people on Sulawesi could have arrived in a similar manner at an equivalent, earlier or later time.

Our findings at Talepu attest to the presence of early tool-makers on Sulawesi by the late Middle Pleistocene, but the absence of Pleistocene human fossils on the island precludes a definitive answer as to which hominin species was first to make landfall. With regards to potential island colonizers, there are at least three candidates in the region: the known and inferred distributions of *H. floresiensis* on Flores (~190 ka (ref. 25) or earlier¹), *H. erectus* on the southern margin of Sunda (present-day Java) (~1.5 million years ago to ~140 ka (refs 26, 27)), and 'Denisovans', whose geographic range may have extended into Wallacea²⁸. Considering the predominantly southerly flowing currents of the Indonesian through-flow²⁹, we speculate that the most likely points of origin for the Sulawesi colonizers are Borneo to the west (part of mainland Asia during periods of low sea level) and the Philippines to the north³⁰ (the northern extremity of Wallacea), with the implication being that other islands in the region harbour undiscovered records of archaic hominins.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 April; accepted 13 November 2015.

1. Brumm, A. et al. Hominins on Flores, Indonesia, by one million years ago. *Nature* **464**, 748–752 (2010).
2. Clarkson, C. et al. The archaeology, chronology and stratigraphy of Madjedbebe (Malakunanja II): a site in northern Australia with early occupation. *J. Hum. Evol.* **83**, 46–64 (2015).
3. O'Connor, S. in *Emergence and Diversity of Modern Human Behaviour in Paleolithic Asia* (eds Yousuke Kaifu et al.) 214–224 (Texas A&M Univ. Press, 2015).
4. Morwood, M. & Van Oosterzee, P. *The Discovery of the Hobbit: The Scientific Breakthrough that Changed the Face of Human History* (Random House Australia, 2007).
5. Aubert, M. et al. Pleistocene cave art from Sulawesi, Indonesia. *Nature* **514**, 223–227 (2014).
6. Van Heekeren, H. R. Early man and fossil vertebrates on the island of Celebes. *Nature* **163**, 492 (1949).
7. Bartstra, G.-J., Keates, S. G., Basoeki & Kallupa, B. On the dispersion of *Homo sapiens* in eastern Indonesia: the Palaeolithic of south Sulawesi. *Curr. Anthropol.* **32**, 317–321 (1991).
8. van Heekeren, H. R. *The Stone Age of Indonesia* (Martinus Nijhoff, 1972).

9. Hooijer, D. A. Pleistocene vertebrates from Celebes: *Celebochoerus heekereni* nov. gen. nov. spec. *Proc. K. Ned. Akad. Wet.* **51**, 1024–1032 (1948).
10. Hooijer, D. A. Quaternary mammals west and east of Wallace's Line. *Neth. J. Zool.* **25**, 46–56 (1974).
11. van den Bergh, G. D. The Late Neogene elephantoid-bearing faunas of Indonesia and their palaeozoogeographic implications: a study of the terrestrial faunal succession of Sulawesi, Flores and Java, including evidence for early hominid dispersal east of Wallace's Line. *Scripta Geologica* **117**, 1–419 (1999).
12. Grainger, A. M. & Davies, K. G. Reef exploration in the East Sengkang Basin, Sulawesi. In *Proc. 12th Annual Convention of the Indonesian Petroleum Association*, V1, 207–227 (1983).
13. Sukanto, R. *Geological Map of Indonesia, Sheet VIII: Ujung Pandang*. Scale 1:1,000,000. Geological Research and Development Centre, Bandung (1975).
14. Grün, R., Eggins, S., Kinsley, L., Moseley, H. & Sambridge, M. Laser ablation U-series analysis of fossil bones and teeth. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **416**, 150–167 (2014).
15. Sambridge, M., Grün, R. & Eggins, S. U-series dating of bone in an open system: the diffusion-adsorption-decay model. *Quat. Geochronol.* **9**, 42–53 (2012).
16. Li, B. & Li, S.-H. Luminescence dating of K-feldspar from sediments: a protocol without anomalous fading correction. *Quat. Geochronol.* **6**, 468–479 (2011).
17. Li, B., Jacobs, Z., Roberts, R. G. & Li, S.-H. Review and assessment of the potential of post-IR IRSL dating methods to circumvent the problem of anomalous fading in feldspar luminescence. *Geochronometria* **41**, 178–201 (2014).
18. O'Connor, S., Ono, R. & Clarkson, C. Pelagic fishing at 42,000 years before the present and the maritime skills of modern humans. *Science* **334**, 1117–1121 (2011).
19. Summerhayes, G. R. *et al.* Human adaptation and plant use in highland New Guinea 49,000 to 44,000 years ago. *Science* **330**, 78–81 (2010).
20. Mellars, P. Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. *Proc. Natl Acad. Sci. USA* **103**, 9381–9386 (2006).
21. Storm, P. *et al.* U-series and radiocarbon analyses of human and animal remains from Wajak, Indonesia. *J. Hum. Evol.* **64**, 356–365 (2013).
22. Barker, G. *et al.* The 'human revolution' in lowland tropical Southeast Asia: the antiquity and behavior of anatomically modern humans at Niah Cave (Sarawak, Borneo). *J. Hum. Evol.* **52**, 243–261 (2007).
23. Grün, R. *et al.* U-series and ESR analyses of bones and teeth relating to the human burials from Skhul. *J. Hum. Evol.* **49**, 316–334 (2005).
24. Westaway, K. E. *et al.* Age and biostratigraphic significance of the Punung Rainforest Fauna, East Java, Indonesia, and implications for *Pongo* and *Homo*. *J. Hum. Evol.* **53**, 709–717 (2007).
25. Westaway, K. E. *et al.* Establishing the time of initial human occupation of Liang Bua, western Flores, Indonesia. *Quat. Geochronol.* **2**, 337–343 (2007).
26. Zaim, Y. *et al.* New 1.5 million-year-old *Homo erectus* maxilla from Sangiran (Central Java, Indonesia). *J. Hum. Evol.* **61**, 363–376 (2011).
27. Indriati, E. *et al.* The age of the 20 meter Solo River terrace, Java, Indonesia and the survival of *Homo erectus* in Asia. *PLoS One* **6**, e21562 (2011).
28. Cooper, A. & Stringer, C. B. Paleontology. Did the Denisovans cross Wallace's Line? *Science* **342**, 321–323 (2013).
29. Sprintall, J. *et al.* The Indonesian seas and their role in the coupled ocean–climate system. *Nature Geosci.* **7**, 487–492 (2014).
30. Mijares, A. S. *et al.* New evidence for a 67,000-year-old human presence at Callao Cave, Luzon, Philippines. *J. Hum. Evol.* **59**, 123–132 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported by grants from the Australian Research Council (ARC) to M.J.M. (DP0770234/DP1093342) and G.D.v.d.B. (FT100100384), and funds from the Geological Survey Institute of Indonesia. L.B. and R.G.R. are supported by ARC Fellowships FT140100384 and FL130100116, respectively. A.B.'s involvement was supported by ARC fellowship DE130101560. The stone tool analysis was supported by ARC Fellowship DP1093342 to M.W.M., and M.S. was funded by the Villum Foundation. The fieldwork was authorized by the directors of the Geological Survey Institute of Indonesia, A. Djumarma Wirakusumah and Y. Kusumahbrata. We further acknowledge the Indonesian State Ministry of Research and Technology and the National Centre for Archaeology in Jakarta (ARKENAS), for facilitating the research. Field assistants included H. Oktaviana, Dadang, S. Sudjarwadi, Ngiliman, T. Suryana and U. P. Wibowo. During the excavations we were assisted by the landowner, Wahe, and 11 other local labourers. Y. Jafari prepared the sediment samples for optical dating and L. Kinsley assisted with the uranium-series analysis. We thank K. Westaway, D. Granger, B. Pillans and B. Jones for additional field support, and S. van der Mije for allowing access to the vertebrate collection at Naturalis. S. Hayes is thanked for providing feedback on the manuscript.

Author Contributions M.J.M. and G.D.v.d.B. conceived the study with F.A., as part of a wider project led by M.J.M., in collaboration with A.B., I.K., S. and E.S. Samples for optical dating were collected and analysed by B.L. and R.G.R. R.G. conducted the uranium-series dating and M.S. analysed samples for $^{40}\text{Ar}/^{39}\text{Ar}$ dating. A.B. and M.W.M. identified and analysed the stone artefacts. G.D.v.d.B. and I.K. analysed the fossil specimens. G.D.v.d.B. and R.S. recorded the site stratigraphy. D.Y. collected and analysed samples for palaeomagnetism. S. conducted a regional geological survey supervised by G.D.v.d.B. and M.J.M. G.D.v.d.B. and A.B. wrote the manuscript, with contributions from the other authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.D.v.d.B. (gert@uow.edu.au) or B.L. (bli@uow.edu.au; about the optical dating method).

METHODS

Excavation methods. In 2007 and 2008 we undertook three (T1–T3) 1 m × 2 m test excavations at Talepu Hill, where large numbers of stone artefacts were found scattered on the surface with loose gravel. The summit of Talepu Hill (4° 22' 06.5'' S; 119° 59' 01.7'' E) lies 36 m above sea level and 18 m above the floodplain of the Walanae River, which flows 600 m to the east (Extended Data Fig. 1). Geological outcrop conditions are very poor, and thick tropical soils cover the underlying geological formations. The three test excavations near the summit of Talepu Hill proved the occurrence of *in situ* stone artefacts down to a depth of at least 1.8 m, in heavily weathered conglomerate lenses and sandy silt layers. The same gravel unit occurs on other hilltops to the west and southwest. At Bulu Palece, 850 m west of Talepu Hill, which is the highest hilltop in the vicinity with an elevation of 51 m (see Extended Data Fig. 2), the gravel is at least 13 m thick, but at Talepu Hill only a basal interval of 4.3 m thickness remains.

In October 2009, T2 was taken down to 7 m below surface (Extended Data Fig. 2b), at which depth the excavation area was reduced to a 1 m × 1 m square and taken down further to a maximum depth of 10 m. To ensure that this deep-trench operation was undertaken safely, we installed timber shoring as the work progressed (Extended Data Fig. 2c). A new east–west oriented, 1 m × 9 m trench (T4) was excavated at the base of the Talepu Hill, 40 m east of T2. This trench reached a maximum depth of 2 m, revealing the lateral development of the stratigraphy near the base of the hill (Fig. 2 and Extended Data Fig. 2d). Deposits were removed in 10 cm spits within stratigraphic units. Stone artefacts and fossils found by the excavators were bagged and labelled immediately; all other deposits were dry sieved with 5 mm mesh to separate out clasts, including stone artefacts. Pebbles from each spit were weighed; and composition analysis was undertaken on clasts from a representative sample from six spits: average maximum clast diameter was recorded by measuring the longest diameter of the ten largest clasts per spit (Extended Data Fig. 3). Bulk samples of stratigraphic units were taken for sediment and pollen analyses.

In October 2010, the excavations at Talepu were continued. A 1 m × 2 m area at the east end of T4 was excavated to a depth of 6.20 m below the surface, thus providing an additional 6 m stratigraphically below the section covered by excavation 2 in 2009. The T4 deposits were removed in 20 cm spits within stratigraphic units. After the excavation of an *in situ* stone artefact (specimen S-TLP10-1, a flake from sub-unit E₂ at a depth of 2.38 m below the surface) and fossils of *Celebochoerus*, it was decided to wet-sieve all the excavated sediments with 3 mm mesh to separate out stones and other clasts, including stone artefacts. Wet-sieving of the silty clay deposits from the interval between 2 and 2.4 m depth yielded one more stone artefact (S-TLP10-2; Fig. 3m) and two possible stone artefacts (S-TLP10-3 and S-TLP10-4). Magnetic susceptibility measurements were taken from the excavation profile at 1 cm intervals with a Bartington MS-2 device, to examine the presence of cryptic tephra layers suitable for dating. A sample for ⁴⁰Ar/³⁹Ar dating was taken at 2.5 m below the ground surface from an interval with elevated magnetic susceptibility values.

In October 2012, backfill of T2 and T4 was removed. T4 was enlarged with a 1 m × 2 m extension (T4-B), and both T2 and T4/4-B were taken further down with an additional 2 m and 2.1 m, respectively, to allow for sampling for palaeomagnetic and optical dating methods. In T4-B two more stone artefacts (Fig. 3j–k), originating from spit 31 (depth 3.0–3.1 m depth below ground level), were recovered on the sieves.

Stone artefacts were analysed following the definitions and methods in ref. 31. The analysis focused on stone-flaking techniques, sequences of reduction, and sizes of stone-flaking products and by-products (Supplementary Tables 1 and 2). The stone artefacts are stored at the Geology Museum in Bandung.

Uranium-series dating. The details for laser ablation uranium-series analysis of skeletal materials were recently summarized¹⁴. Uranium-series analyses provide insights into when uranium migrates into a bone or tooth. This may happen a short time after the burial of the skeletal element or some significant time span later. There may also be later uranium-overprints that are difficult to recognize. As such, apparent uranium-series results from faunal remains have generally to be regarded as minimum age estimates. It is very difficult or impossible to evaluate by how much the uranium-series results underestimate the correct age of the sample. Details of the instrumentation, analytical procedures and data evaluation have been modified from those described in detail elsewhere^{14,32}. All isotope ratios refer to activity ratios.

Sequential laser spot analyses were undertaken on cross sections of eight *Celebochoerus* fossils from the T4 excavation at Talepu. They comprised fragments of six teeth and two bones from sub-unit E₂ found 10–50 cm below the lowest stone artefacts in the same silt layer. Of one fossil (TLP10-1, a *Celebochoerus* lower canine), two subsamples were analysed (a and b). Each fossil specimen was cut transversely using a dentist drill with a diamond saw blade (Extended Data Fig. 4).

Four or five samples were then mounted together into aluminium cups, aligning the cross-sections with the outer rim of the sample holder, which later positioned the samples on the focal plane of the laser. Uranium-series isotopes were measured using the laser ablation multicollector (MC)-ICP-MS system at The Australian National University's (ANU) Research School of Earth Sciences. It consists of a Finnigan MAT Neptune MC-ICP-MS equipped with multiple Faraday cups. At the time of measurement, the mass spectrometer had only one ion counter. This necessitated two sequential sets of measurements along parallel tracks, one for ²³⁰Th and a second for ²³⁴U. The ion counter was set either to masses 230.1 or 234.1 while the Faraday cups measured the masses 232, 235 and 238. Samples were ablated with a Lambda Physik LPFPro ArF excimer ($\lambda = 193$ nm) laser coupled to the Neptune through an ANU-designed Helex ablation cell.

The samples were initially cleaned for 10 s with the laser spot size set to 265 μ m followed by a 50 s analysis run with a 205 μ m spot size using a 5 Hz pulse rate. Analyses were performed at regular intervals along traverses, all starting from the exterior surface (Extended Data Fig. 4a–i). The data sets of each transect were bracketed between reference standard analyses to correct for instrument drift.

Semi-quantitative analysis of uranium and thorium concentrations were derived from repeated measurements of the SRM NIST-610 glass (uranium = 461.5 μ g g⁻¹; thorium = 457.2 μ g g⁻¹), and uranium-isotope ratios from repeated measurements of rhinoceros tooth dentine from Hexian (sample 1118)³³. Age estimates combining all measurements on a specimen were calculated using the iDAD program¹⁵, assuming diffusion from both surfaces for the bones (TLP10-6 and 7) and roots of the teeth (Extended Data Fig. 4a–f, h, i) and directional diffusion from the central pulp cavity into the dentine and covering enamel for TLP10-9 (Extended Data Fig. 4g). The enamel data of the enamel samples were omitted as enamel has a different diffusion rate. Generally, results with elemental U/Th < 300 are rejected, as these are associated with detrital contamination. However, this applied only to a single measurement. The finite ages are given with 2 σ error bands; the infinite results only refer to the lower bound of the 2 σ confidence interval (Supplementary Table 3). None of the samples showed any indication for uranium leaching, which is either expressed by sections with ²³⁰Th/²³⁴U >> ²³⁴U/²³⁸U or increasing ²³⁰Th/²³⁴U ratios towards the surface in conjunction with decreasing uranium-concentrations.

Five samples had infinite positive error bounds and it was thus only possible to calculate minimum ages. It can be seen that the uranium-series results may change over small distances within a sample. The first data set of TLP10-1 yielded a finite result of 161 ± 15 kyr while the second set yielded a minimum age of >255 kyr. As mentioned above, all uranium-series results, whether they are finite or infinite, have to be regarded as minimum age estimates. If the faunal elements present a single population, the uranium-series results indicate that the Talepu samples are most probably older than ~350 kyr, but certainly older than ~200 kyr (Supplementary Table 3). The large errors do not allow us to further constrain the age.

Infrared stimulated luminescence dating of feldspar grains. Optical dating provides an estimate of the time since grains of quartz or potassium-rich feldspar were last exposed to sunlight^{34–37}. The burial age is estimated by dividing the equivalent dose (D_e , a measure of the radiation energy absorbed by grains during their period of burial) by the environmental dose rate (the rate of supply of ionizing radiation to the grains over the same period). D_e is determined from the laboratory measurements of the optically stimulated luminescence (OSL) from quartz or the infrared stimulated luminescence (IRSL) from potassium (K)-feldspar, and the dose rate is estimated from laboratory and field measurements of the environmental radioactivity.

K-feldspar has two advantages over quartz for optical dating: (1) the IRSL signal (per unit absorbed dose) is usually much brighter than the OSL signal from quartz; and (2) the IRSL traps saturate at a much higher dose than do the OSL traps, which makes it possible to date older samples using feldspars than is feasible using the OSL signal from quartz. However, the routine dating of K-feldspars using the IRSL signal has been hampered by the malign phenomenon of 'anomalous fading' (that is, the leakage of electrons from IRSL traps at a faster rate than expected from kinetic considerations³⁸), which gives rise to substantial underestimates of age unless an appropriate correction is made³⁹. Recently, IRSL traps that are less prone to fading have been identified⁴⁰, using either a post-infrared IRSL (pIRIR) approach^{41,42} or a MET-pIRIR procedure^{16,43}. The progress, potential and remaining problems in using these pIRIR signals for dating have been reviewed recently¹⁷.

Dating the samples from Talepu using quartz OSL is impractical because of the paucity of quartz. Furthermore, the quartz OSL traps are expected to be in saturation, owing to the ages of the samples (>100 kyr) and the high environmental dose rates of the deposits (4–5 Gy/kyr). In this study, we applied the MET-pIRIR procedure to K-feldspar extracts from Talepu to isolate the light-sensitive IRSL signal that is least prone to anomalous fading. We also allowed for any residual dose at the time of sediment deposition, to account for the fact that pIRIR traps are less easily bleached than the 'fast' component OSL traps in quartz. The resulting

MET-pIRIR ages should, therefore, be reliable estimates of the time of sediment deposition at Talepu.

The total environmental dose rate for K-feldspar grains consists of four components: the external gamma, beta and cosmic-ray dose rates, and the internal beta dose rate. The dosimetry data for all samples are summarized in Supplementary Table 4.

The external gamma dose rates were measured using an Exploranium GR-320 portable gamma-ray spectrometer, equipped with a 3-inch diameter NaI(Tl) crystal calibrated for uranium, thorium and potassium concentrations using the CSIRO facility at North Ryde⁴⁴. At each sample location, three or four measurements of 300 s duration were made of the gamma dose rate at field water content. The external beta dose rate was measured by low-level beta counting using a Riso GM-25-5 multiscaler system⁴⁵ and referenced to the Nussloch Loess (Nussi) standard⁴⁶. The external beta dose rate was corrected for the effect of grain size and hydrofluoric acid etching on beta-dose attenuation. These external components of the total dose rate were adjusted for assumed long-term water contents of 20% for the Talepu Upper Trench (TUT = T2) samples and 30% for the Talepu Lower Trench (TLT = T4) sample (TUT and TLT sample numbers refer to the Centre for Archaeological Science laboratory numbers). These values are based on the measured field water contents (Supplementary Table 4), together with an assigned 1σ uncertainty of $\pm 5\%$ to capture the likely range of time-averaged mean values over the entire period of sample burial.

To check the equilibrium status of the ^{238}U and ^{232}Th decay chains, each sample was dried, ground to a fine powder and then analysed by high-resolution gamma-ray spectrometry (HRGS). The measured activities of ^{238}U , ^{226}Ra and ^{210}Pb in the ^{238}U series, ^{228}Ra and ^{228}Th in the ^{232}Th series, and ^{40}K are listed in Supplementary Table 5. The activities of ^{228}Ra and ^{228}Th were close to equilibrium for all of the samples, as is commonly the case with the ^{232}Th series. By contrast, the ^{238}U chain of each sample, except TUT-OSL9, was in disequilibrium at the present day. Sample TUT-OSL2 had a 39–45% deficit of ^{226}Ra and ^{210}Pb relative to the parental ^{238}U activity, whereas sample TUT-OSL3 had a 224–345% excess of the daughter nuclides. Samples TUT-OSL1 and TLT-OSL6 had ^{226}Ra deficits of 50% and 26%, respectively, relative to their ^{238}U activities, but the ^{210}Pb activities of both samples were similar to their parental ^{238}U activities.

Sample TUT-OSL3 was the only sample with a present-day excess of ^{226}Ra . This sample was from a sandy layer (unit B) through which ground water could percolate, so we attributed the observed ^{226}Ra excess to the deposition of radium transported by ground water. Given the similar ^{238}U activities of TUT-OSL3 and nearby TUT-OSL2, it is reasonable to assume that the parental uranium activity had not changed substantially during the period of burial of either sample, and that the ^{226}Ra excess in TUT-OSL3 most probably occurred recently. The latter can be deduced from the fact that ^{226}Ra has a half-life of $\sim 1,600$ years, which is short relative to the ages of our samples (> 100 kyr), so any unsupported excess of ^{226}Ra would have decayed back into equilibrium with ^{238}U within ~ 8 kyr of deposition (that is, five half-lives of ^{226}Ra). The alternative option—that groundwater has continuously supplied excess ^{226}Ra to unit B—is not supported by the disequilibrium between ^{226}Ra and ^{210}Pb : the latter nuclide has a half-life of ~ 22 years, so it should remain in equilibrium with ^{226}Ra if the latter is supplied continuously and no radon gas is lost to atmosphere. Moreover, as the return of ^{210}Pb to equilibrium with ^{226}Ra is governed by the half-life of the shorter-lived nuclide, it could be argued that the excess ^{226}Ra was deposited within the past ~ 110 years (five half-lives of ^{210}Pb).

Fortunately, the calculated age of TUT-OSL3 is not especially sensitive to different assumptions about the timing or extent of disequilibrium in the ^{238}U series. The latter accounts for only 28% of the total dose rate estimated from the HRGS data in Supplementary Table 5; this assumes that the present-day nuclide activities have prevailed throughout the period of sample burial. If, instead, as we consider more likely, the observed excess in ^{226}Ra was deposited recently and the ^{238}U decay chain had been in equilibrium for almost all of the period of sample burial, then the ^{238}U series accounted for only 12% of the total dose rate (that is, using activities of $37 \pm 4 \text{ Bq kg}^{-1}$ for ^{238}U , ^{226}Ra and ^{210}Pb). The ages calculated under these two alternative scenarios, using only the HRGS data for estimating external beta and gamma dose rates, range from ~ 118 kyr to ~ 143 kyr (Supplementary Table 5).

Sample TUT-OSL2 was from the more silty overlying layer (sub-unit A7) and had deficits of ^{226}Ra and ^{210}Pb relative to ^{238}U , but these disequilibria were much smaller in magnitude than those of TUT-OSL3. If it were not continuously leached from the sample, ^{226}Ra will return to secular equilibrium with ^{238}U within ~ 8 kyr, so the existence of disequilibrium in TUT-OSL2 adds further weight to the argument for recent transport of ^{226}Ra in ground water at Talepu. The alternative is that ^{226}Ra has been leached continuously from this sample, so we performed the same sensitivity test on the dose rates and ages as that performed on TUT-OSL3. For TUT-OSL2, the ages determined using the present-day HRGS data or activities of $41 \pm 3 \text{ Bq kg}^{-1}$ for ^{238}U , ^{226}Ra and ^{210}Pb are statistically indistinguishable (130 ± 12

and 125 ± 11 kyr, respectively; Supplementary Table 5), because the disequilibria are much less marked than in TUT-OSL3 and the ^{238}U series makes only a small contribution (10–14%) to the total dose rate of TUT-OSL2.

Samples TUT-OSL1 and TLT-OSL6 had deficits of ^{226}Ra relative to ^{238}U , but similar activities of ^{238}U and ^{210}Pb . The latter additionally strengthens our proposition that ^{226}Ra was leached from these sediments recently, because ^{210}Pb should return to a state of equilibrium with ^{226}Ra within ~ 110 years (five half-lives of ^{210}Pb). For both samples, the ages calculated using the present-day HRGS data were statistically concordant with those estimated by assuming that the ^{238}U chain had been in secular equilibrium for almost the entire period of sample burial (Supplementary Table 5). The same applies to sample TUT-OSL9, since the measured activities of ^{238}U , ^{226}Ra and ^{210}Pb were consistent at 1σ .

To calculate the ages of the Talepu samples, we used the beta dose rates deduced from direct beta counting and the *in situ* gamma dose rates measured at each sample location. The external beta dose rates determined from beta counting and from the HRGS data (Supplementary Table 5) were statistically consistent (at 2σ) for all five samples; such agreement is expected, as both measure the present-day activities. The field gamma dose rates are also based on the nuclide activities prevailing at the time of measurement (^{214}Bi , a short-lived nuclide between ^{226}Ra and ^{210}Pb , being used for the ^{238}U series) and—importantly—take into account any spatial heterogeneity in dose rate from the ~ 30 cm of deposit surrounding each sample.

The *in situ* gamma dose rates for samples TUT-OSL1 and TLT-OSL6 were consistent at 1σ with those estimated from the HRGS activities, whereas the field gamma dose rates for TUT-OSL2, -OSL3 and -OSL9 were either higher or lower than those calculated from the HRGS data. The lower *in situ* gamma dose rate of TUT-OSL3 can be explained by the location of this sample close to the boundary with the TUT-OSL2 sediments, which have a smaller beta dose rate (Supplementary Table 4), and vice versa for the elevated field gamma dose rate of the latter sample. This result also indicates that the ^{226}Ra and ^{210}Pb deficits (TUT-OSL2) and excesses (TUT-OSL3) were spatially localized and not pervasive in the 30 cm of deposit surrounding these samples.

Under dim red laboratory illumination, the collected samples (see Methods) were treated with hydrochloric acid and hydrogen peroxide solutions to remove carbonates and organic matter, then dried. Grains of 90–180 or 180–212 μm in diameter were obtained by dry sieving. The K-feldspar grains were separated from quartz and heavy minerals using a sodium polytungstate solution of density 2.58 g cm^{-3} , and etched in 10% hydrofluoric acid for 40 min to clean the surfaces of the grains and remove (or greatly reduce in volume) the external alpha-irradiated layer of each grain. For each sample, 8–14 aliquots were prepared by mounting grains as a 5-mm-diameter monolayer in the centre of a 9.8-mm-diameter stainless steel disc, using ‘Silkospay’ silicone oil as the adhesive. This resulted in each aliquot consisting of several hundred K-feldspar grains.

The single-aliquot regenerative-dose (SAR) MET-pIRIR procedure introduced in ref. 16 was adapted for the Talepu samples in this study. We modified the original procedure by using a preheat at 320°C (rather than 300°C) for 60 s, to avoid significant influence from residual phosphorescence while recording the MET-pIRIR signal at 250°C (Supplementary Table 6). In addition, following ref. 47, we used a 2 h solar simulator bleach before each regenerative dose cycle, instead of the high-temperature infrared bleaching step used originally, as this proved essential for recovering a given laboratory dose (see below).

Example IRSI (50°C) and MET-pIRIR (100 – 250°C) decay curves are shown in Extended Data Fig. 6a for an aliquot of sample TUT-OSL2. The decay curves observed at the different stimulation temperatures are similar in shape, with initial MET-pIRIR signal intensities of the order of a few thousand counts per second. Extended Data Fig. 6b shows the corresponding dose–response curves for the same aliquot. Each sensitivity-corrected (L_x/T_x) dose–response curve was fitted using a single saturating-exponential function of the form $I = I_0(1 - \exp^{-D/I_0})$, where I is the L_x/T_x value at regenerative dose D , I_0 is the saturation value of the exponential curve and D_0 is the characteristic saturation dose. The D_0 values are shown next to each dose–response curve in Extended Data Fig. 6b. For a total of 38 aliquots drawn from all 5 samples, we calculated the D_0 values for the 250°C MET-pIRIR signal; these are plotted in Extended Data Fig. 6c. On a ‘radial plot’ such as this, the most precise estimates fall to the right and the least precise to the left. If these independent estimates are statistically consistent with a common value at 2σ , then 95% of the points should scatter within a band of width ± 2 units projecting from the left-hand (‘standardized estimate’) axis to the common value on the right-hand, radial axis. The radial plot thus provides simultaneous information about the spread, precision and statistical consistency of experimental data^{48–50}. The measured D_0 values range from ~ 220 to ~ 600 Gy, with the vast majority consistent at 2σ with a common value of ~ 360 Gy. The average D_0 value (calculated using the central age model⁴⁹) is 358 ± 14 Gy, with the standard error taking the extent of overdispersion ($16 \pm 4\%$) into account. If we

adopt the D_0 values corresponding to 90% ($2.3D_0$) and 95% ($3D_0$) of the saturation level of the typical dose–response curve as the upper limits for reliable estimation of D_e ^{43,47,51,37}, then the maximum reliable D_e values that we can determine using the 250 °C MET-pIRIR signal are ~820 Gy and ~1070 Gy, respectively, for these samples.

To validate whether the MET-pIRIR procedure is applicable to the Talepu samples, we conducted dose recovery, anomalous fading and residual dose tests. For the latter, four aliquots of each sample were bleached for 4–5 h using a Dr Hönle solar simulator (model UVACUBE 400). The residual doses were then estimated by measuring these bleached aliquots using the modified MET-pIRIR procedure (Supplementary Table 6). The residual doses obtained for each of the TUT samples are plotted against stimulation temperature in Extended Data Fig. 7a. The IRSL signal measured at 50 °C has a few grays of residual dose, which increases as the stimulation temperature is raised, attaining values of 16–20 Gy at 250 °C. The size of the residual dose is only about 2–3% of the corresponding D_e values for the 250 °C signal, which were subtracted from the D_e values for the respective samples before calculating their ages.

It was noted in ref. 52 that a simple subtraction of the residual dose from the apparent D_e value could result in underestimation of the true D_e value if the residual signal is large relative to the bleachable signal. Accordingly, it advocated the use of an ‘intensity-subtraction’ procedure instead of the simple ‘dose-subtraction’ approach for samples with large residual doses. The dose-subtraction approach should be satisfactory for the Talepu samples, however, given the small size of the residual doses compared with the D_e values obtained from the MET-pIRIR 250 °C signal.

A dose recovery test⁴⁹ was conducted on sample TUT-OSL1. Eight aliquots were bleached by the solar simulator for 5 h, then given a ‘surrogate natural’ dose of 550 Gy. Four of these aliquots were measured using the original MET-pIRIR procedure¹⁶, with a ‘hot’ infrared bleach of 320 °C for 100 s applied at the end of each SAR cycle (step 15 in Supplementary Table 6). The other four aliquots were measured using the modified MET-pIRIR procedure (Supplementary Table 6), with a solar simulator bleach of 2 h used at step 15. The measured doses at each stimulation temperature were then corrected for the corresponding residual doses (Extended Data Fig. 7a), and the ratios of measured dose to given dose were calculated for the IRSL and MET-pIRIR signals.

The dose recovery ratios are plotted in Extended Data Fig. 7b, which shows that a hot bleach at the end of each SAR cycle results in significant overestimation of the known (given) dose; for the MET-pIRIR 250 °C signal, an overestimation of 48% was observed. For these same four aliquots, we obtained a ‘recycling ratio’ (the ratio of the L_x/T_x signals for two duplicate regenerative doses) consistent with unity (1.00 ± 0.03), which indicates that the test-dose sensitivity correction worked successfully between regenerative-dose cycles. The overestimation in recovered dose, therefore, implies failure of the sensitivity correction for the surrogate natural dose: that is, the extent of sensitivity change between measurement of the surrogate natural and its corresponding test dose differs from the changes occurring in the subsequent regenerative-dose cycles. The surrogate natural and regenerative-dose cycles differ only in respect to the preceding bleaching treatment (that is, a solar simulator bleach was used for the former and a hot bleach for the latter), so we compared these results with those obtained for the four aliquots that were bleached at the end of each regenerative-dose cycle using the solar simulator. The dose recovery results improved significantly using this modified procedure (Extended Data Fig. 7): all of the measured/given dose ratios were consistent with unity (at 2σ) for the signals measured at different temperatures, with a ratio of 1.02 ± 0.03 obtained for the MET-pIRIR 250 °C signal.

The results of the dose recovery test on sample TUT-OSL1 suggest that the MET-pIRIR procedure could successfully recover a known dose given to K-feldspars from Talepu, but only when a solar simulator bleach was applied at the end of each SAR cycle. We therefore adopted this procedure to measure the D_e values for all five Talepu samples.

Previous studies of pIRIR signals have shown that the anomalous fading rate (g value) depends on the stimulation temperature, with negligible fading of MET-pIRIR signals stimulated at temperatures of 200 °C and above^{16,17}. Accordingly, no fading correction is required for these high-temperature MET-pIRIR signals. To check that this finding also applied to the Talepu samples, fading tests were conducted on six aliquots of sample TUT-OSL3 that had already been used for D_e measurements. We adopted a single-aliquot procedure similar to that described in ref. 53, but based on the MET-pIRIR signals. Doses of 110 Gy were administered using the laboratory beta source, and the irradiated aliquots were then preheated and stored for periods of up to 1 week at room temperature (~20 °C). For practical reasons, we used a hot bleach (320 °C for 100 s) instead of a solar simulator bleach at the end of each SAR cycle, but this choice should not have affected the outcome of the fading test, given the aforementioned recycling ratio of

unity obtained using the hot bleach. Extended Data Fig. 7c shows the decay in the sensitivity-corrected MET-pIRIR signal as a function of storage time for these six aliquots, normalized to the time of prompt measurement (which ranged from 720 s for the 50 °C IRSL to 1480 s for the 250 °C MET-pIRIR signal). The corresponding fading rates (g values) were calculated for the IRSL and MET-pIRIR signals (Extended Data Fig. 7d). The highest fading rate was observed for the 50 °C IRSL signal ($5.5 \pm 0.4\%$ per decade), and decreases as the stimulation temperature is increased, falling to 0.94 ± 0.92 and $0.17 \pm 1.13\%$ per decade for the 200 and 250 °C signals, respectively. The latter g value is consistent with zero at 1σ , so we used the D_e value obtained from the 250 °C signal to date each of the samples. We note, however, that the g values for the 200 and 250 °C signals have large uncertainties, owing to the difficulty in obtaining precise estimates at low fading rates, so our data do not exclude the possibility that the high-temperature signals may fade slightly.

On the basis of the results of the performance tests described above, the MET-pIRIR procedure in Supplementary Table 6 was used to estimate the D_e values for all four TUT samples, as well as one sample (TLT-OSL6) collected from near the base of the stratigraphically underlying deposits in the TLT. The D_e estimates obtained for the TUT samples using the MET-pIRIR 250 °C signal are shown in Extended Data Fig. 8. Most of the estimates are distributed around a central value, although the spread is larger than can be explained by the measurement uncertainties alone. The overdispersion among these D_e values is ~20% for three of the TUT samples and almost twice this amount for TUT-OSL9, the latter arising from a pair of low D_e values measured with relatively high precision. To estimate the age for each of these samples, we determined the weighted mean D_e of the individual single-aliquot values using the central age model⁴⁹, which takes account of the measured overdispersion in the associated standard error.

As a further test of the reliability of our D_e estimates for the TUT samples, we have plotted the estimates of the central age model as a function of stimulation temperature in Extended Data Fig. 9a. These plots show that the D_e values increase with stimulation temperature until a ‘plateau’ is reached at higher temperatures for each of the TUT samples; the plateau region (marked by the dashed line) indicates that a non-fading component is present at these elevated temperatures. The existence of a plateau can be used, therefore, as an internal, diagnostic tool to confirm that a stable, non-fading component has been isolated for age determination. For all four TUT samples, a plateau is reached at temperatures of 200 °C and above, from which we infer negligible fading of the MET-pIRIR 250 °C signal. We calculated the sample ages, therefore, using the D_e values obtained from the 250 °C signal. The corresponding weighted mean D_e values, dose rate data and final ages are listed in Supplementary Table 4.

For sample TLT-OSL6 from the TLT, four of the eight aliquots measured emitted natural MET-pIRIR 250 °C signals consistent with the saturation levels of the corresponding dose–response curves (for example Extended Data Fig. 9b). This implies that the IRSL traps were saturated in the natural sample, which further supports our conclusion that the MET-pIRIR 250 °C signal had a negligible fading rate. It would be hazardous to estimate the age of sample TLT-OSL6 from the D_e values of the four non-saturated aliquots, as these may represent only the low D_e values in the ‘tail’ of a truncated distribution. If we adopt the average $2.3D_0$ value for the MET-pIRIR 250 °C signal of all five Talepu samples (~820 Gy) as an upper limit for reliable D_e estimation, then this corresponds to a minimum age of ~195 kyr for sample TLT-OSL6 (Supplementary Table 4).

K-feldspar chronology. The MET-pIRIR 250 °C ages for the four samples dated from the TUT (=T2) are in correct stratigraphic order, increasing from 103 ± 9 kyr (at ~3 m depth) to 156 ± 19 kyr (at ~10 m depth). They thus span the period from marine isotope stage 6—the penultimate glacial—to marine isotope stage 5, the last interglacial. This coherent sequence of ages also supports our contention that the Talepu samples were sufficiently bleached before deposition.

The sample analysed from ~8 m depth in the TLT (=T4; sample TLT-OSL6) yielded a minimum age of ~195 kyr, corresponding to marine isotope stage 7 (the penultimate interglacial) or earlier. We have not yet dated the other sediments exposed in the TLT, but expect that the 6 m of deposit immediately overlying TLT-OSL6 will be older than 156 ± 19 kyr, as they stratigraphically underlie sample TUT-OSL9 in the TUT.

We interpret the ages for the TUT samples as true (finite) depositional ages, based on the existence of D_e plateaux (Extended Data Fig. 9a) and the increase in D_e with depth (that is, ordered stratigraphically). This is the most parsimonious reading of our data. The measured fading rate of $0.17 \pm 1.13\%$ per decade for sample TUT-3 allows for the possibility, however, that the MET-pIRIR 250 °C signal may still fade slightly and that our samples had reached an equilibrium state of trap filling and emptying (so-called field saturation⁵⁴). If so, then the increase in D_e with depth could, instead, be due to a systematic decline in fading rate with increasing depth. Any such a trend cannot be verified or rejected from laboratory measurements of the g value, owing to the size of the associated uncertainties at

low fading rates (Extended Data Fig. 7c, d). The ages for the TUT samples could, therefore, be viewed conservatively as minimum ages (as for sample TLT-OSL6), given the uncertainties in the measured fading rate of the 250 °C signal and the exact level at which the signal saturates. The measured age of the uppermost sample in the sequence, TUT-OSL1, would increase by about 15% and 40% after correcting^{39,55,56} for assumed fading rates of 0.5 and 1% per decade, respectively. Similarly, the measured ages of TUT-OSL2, -3 and -4 would increase by about 17, 23 and 28%, respectively, after correcting for an assumed fading rate of 0.5% per decade. Thus, whether viewed as true ages or as minimum ages, the TUT sediments were deposited more than ~100 ka.

Palaeomagnetic dating. Samples for palaeomagnetic polarity assessment were taken from the baulks of excavations Talepu 2 (T2) and Talepu 4 (T4) (Fig. 2). Samples were taken at 20–30 cm intervals using non-magnetic tools. Preferably samples in non-bioturbated silty deposits were taken. The upper conglomeratic interval of T2 was omitted because of its coarser grain size and because it appeared heavily affected by soil formation and plant root bioturbation. From each sample level, five oriented sample specimens were retrieved by carving the sediment using non-magnetic tools and fitting them into 8 cm³ plastic cubes. The samples were labelled according to excavation, baulk and depth.

In the laboratory all specimens were treated by an alternating field demagnetizer. The mean magnetic directions for each sample are presented in Supplementary Table 7. Demagnetization was performed with intervals of 2.5–5 mT to a peak of up to 80–1,000 mT. The magnetization vectors obtained from most samples showed no more than two separated components of natural remanent magnetization (NRM) on the orthogonal planes, which means that the specimens had been affected by secondary magnetization. However, secondary magnetization was easily removed with a demagnetization of up to 5–20 mT, while the characteristic remanent magnetizations (ChRMs) could be isolated through stepwise demagnetization of up to 20–40 mT, in some cases up to 50 mT. Above 40 mT most samples were completely demagnetized (Extended data Fig. 4j and Supplementary Table 8).

The mean magnetization intensities and palaeomagnetic directions are plotted against stratigraphic depth in Extended Data Fig. 5. The 90–98% intensity saturation was achieved from 1.30×10^{-4} to 3.81×10^{-3} A m⁻¹ before demagnetization, and between 8.52×10^{-6} and 1.49×10^{-4} A m⁻¹ after demagnetization at 20–40 mT. The direction of ChRMs is determined from the orthogonal plots in at least four or five successive measurement steps between 20 and 50 mT using principal component analysis⁵⁷ (PuffinPlot⁵⁸ and IAPD 2000 software⁵⁹) with the maximum angular deviations setting at <5°. Although there are no well-defined criteria for the acceptability of palaeomagnetic data available, the $k > 30$ and $\alpha_{95} < 15^\circ$ criteria of ref. 60 were used to accept the average remanence direction for sampled levels. On the basis of these tests, all the samples ($n = 24$) throughout the Talepu sequences yielded acceptable ChRM directions and showed a normal polarity. The ChRM directions were relatively constant throughout the sequences, except the direction of samples taken in T2 at 6.5 and 7.5 m depth, which showed steep inclinations of 56–68°. Such steep inclinations are unusual for near-equatorial regions. One possible interpretation is that post-depositional mass-movement disturbances, such as creep or a landslide, resulted in rotational movements of this interval.

The equal-area projections show that the dispersion of within-site means of the remanence directions re-group more closely together after demagnetization, and no significant change in the major remanence direction occurs with depth. The major remanent direction corresponds closely with the present magnetization direction (Extended Data Fig. 5b).

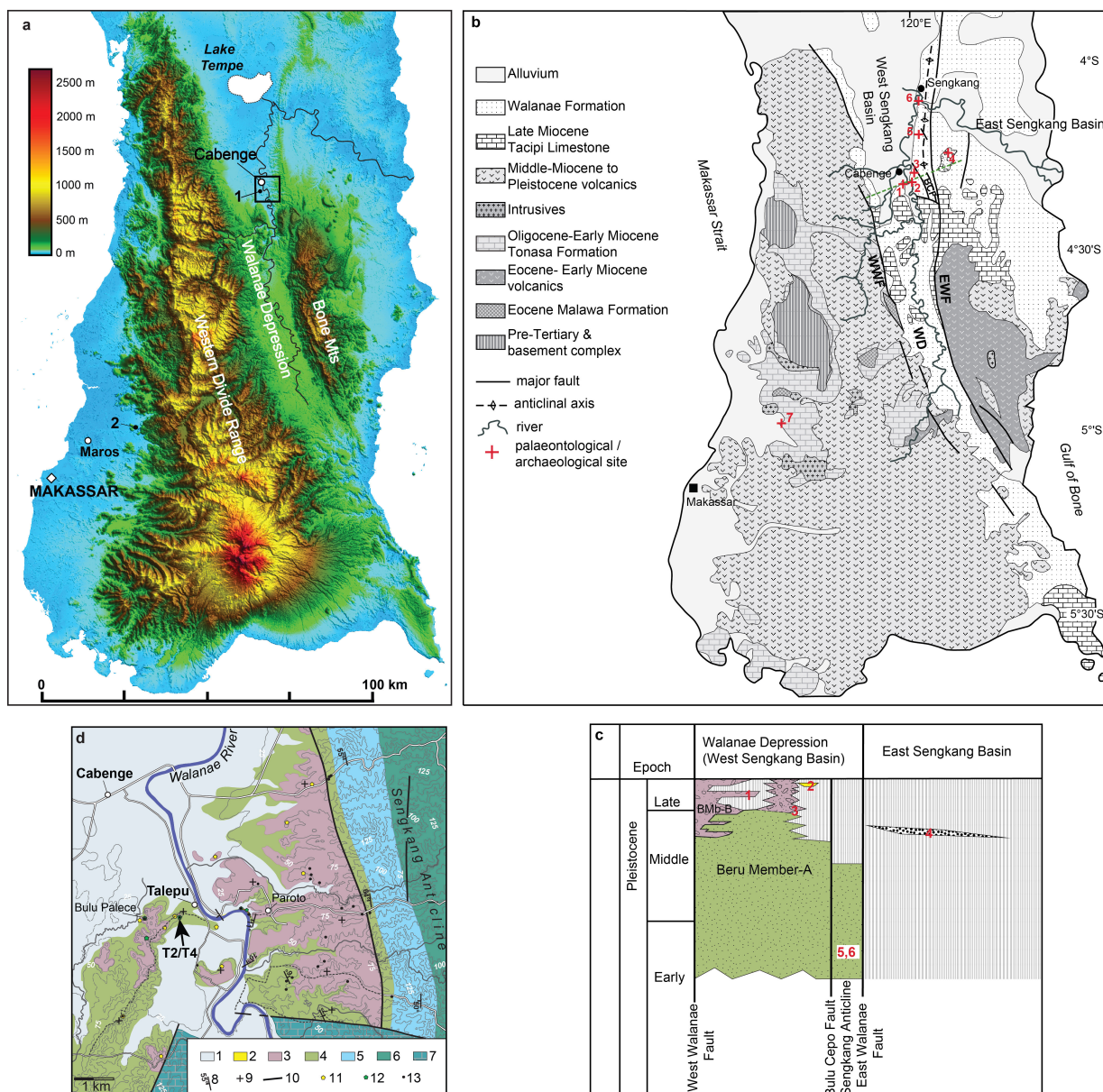
⁴⁰Ar/³⁹Ar fusion dating of single sanidine crystals. Sample TAL-10-01 was taken from T4, sub-unit E₂, at a depth of 2.5 m below the surface. Euhedral sanidine crystals up to 250 µm in length were hand-picked following standard heavy liquid and magnetic separation techniques. Crystals were loaded into wells in aluminium sample discs (diameter 18 mm) for neutron irradiation, along with the 1.185 Myr Alder Creek sanidine⁶¹ as the neutron fluence monitor. Neutron irradiation was done in the cadmium-shielded CLICIT facility at the Oregon State University TRIGA reactor. Argon isotopic analyses of gas released by CO₂ laser fusion of single sanidine crystals (Supplementary Table 9) were made on a fully automated, high-resolution, Nu Instruments Noblesse multi-collector noble-gas mass spectrometer, using procedures documented previously^{1,62}. Sample gas clean-up was through an all-metal extraction line, equipped with a -130 °C cold trap (to remove H₂O) and two water-cooled SAES GP-50 getter pumps (to absorb reactive gases). Argon isotopic analyses of unknowns, blanks and monitor minerals were performed in identical fashion. ⁴⁰Ar and ³⁹Ar were measured on the high-mass ion counter, ³⁸Ar and ³⁷Ar on the axial ion counter and ³⁶Ar on the low-mass ion counter, with baselines measured no less than every third cycle. Measurement of the ⁴⁰Ar, ³⁸Ar and ³⁶Ar ion beams was performed simultaneously, followed by

sequential measurement of ³⁹Ar and ³⁷Ar. Beam switching was achieved by varying the field of the mass spectrometer magnet and with minor adjustment of the quad lenses. Data acquisition and reduction was performed using the program 'Mass Spec' (A. Deino, Berkeley Geochronology Center). Detector intercalibration and mass fractionation corrections were made using the weighted mean of a time series of measured atmospheric argon aliquots delivered from a calibrated air pipette. Decay and other constants, including correction factors for interference isotopes produced by nucleogenic reactions, are as reported in ref. 62.

The resulting age probability diagram for single sanidine crystals (Extended Data Figure 10) shows a wide range in ages with a dominant population around 9.4 million years ago (Late Miocene). This indicates that the sanidine crystals from the sample do not represent a single volcanic event, but were predominantly derived from erosion of the Miocene volcanic rocks west of the Wanae Depression and/or from Late Miocene marine sediments of the Wanae Formation.

- Moore, M. W., Sutikna, T., Jatmiko, Morwood, M. J. & Brumm, A. Continuities in stone flaking technology at Liang Bua, Flores, Indonesia. *J. Hum. Evol.* **57**, 503–526 (2009).
- Eggins, S. M. *et al.* In situ U-series dating by laser-ablation multi-collector ICPMS: new prospects for Quaternary geochronology. *Quat. Sci. Rev.* **24**, 2523–2538 (2005).
- Grün, R. *et al.* ESR and U-series analyses of teeth from the palaeoanthropological site of Hexian, Anhui Province, China. *J. Hum. Evol.* **34**, 555–564 (1998).
- Huntley, D. J., Godfrey-Smith, D. I. & Thewalt, M. L. Optical dating of sediments. *Nature* **313**, 105–107 (1985).
- Aitken, M. J. *An Introduction to Optical Dating: The Dating of Quaternary sediments by the Use of Photon-Stimulated Luminescence* (Oxford Univ. Press, 1998).
- Lian, O. B. & Roberts, R. G. Dating the Quaternary: progress in luminescence dating of sediments. *Quat. Sci. Rev.* **25**, 2449–2468 (2006).
- Roberts, R. G. *et al.* Optical dating in archaeology: thirty years in retrospect and grand challenges for the future. *J. Archaeol. Sci.* **56**, 41–60 (2015).
- Wintle, A. G. Anomalous fading of thermo-luminescence in mineral samples. *Nature* **245**, 143–144 (1973).
- Huntley, D. J. & Lamothe, M. Ubiquity of anomalous fading in K-feldspars and the measurement and correction for it in optical dating. *Can. J. Earth Sci.* **38**, 1093–1106 (2001).
- Thomsen, K. J., Murray, A., Jain, M. & Bøtter-Jensen, L. Laboratory fading rates of various luminescence signals from feldspar-rich sediment extracts. *Radiat. Meas.* **43**, 1474–1486 (2008).
- Buylaert, J.-P., Murray, A., Thomsen, K. J. & Jain, M. Testing the potential of an elevated temperature IRSL signal from K-feldspar. *Radiat. Meas.* **44**, 560–565 (2009).
- Thiel, C. *et al.* Luminescence dating of the Stratzing loess profile (Austria)—testing the potential of an elevated temperature post-IR IRSL protocol. *Quat. Int.* **234**, 23–31 (2011).
- Li, B. & Li, S.-H. Luminescence dating of Chinese loess beyond 130 ka using the non-fading signal from K-feldspar. *Quat. Geochronol.* **10**, 24–31 (2012).
- Hutton, J. & Prescott, J. Field and laboratory measurements of low-level thorium, uranium and potassium. *Nucl. Tracks Radiat. Meas.* **20**, 367–370 (1992).
- Bøtter-Jensen, L. & Mejdahl, V. Assessment of beta dose-rate using a GM multi-counter system. *Nucl. Tracks Radiat. Meas.* **14**, 187–191 (1988).
- Jacobs, Z. & Roberts, R. G. An improved single grain OSL chronology for the sedimentary deposits from Diepkloof Rockshelter, Western Cape, South Africa. *J. Archaeol. Sci.* **63**, 175–192 (2015).
- Li, B., Roberts, R. G., Jacobs, Z. & Li, S.-H. A single-aliquot luminescence dating procedure for K-feldspar based on the dose-dependent MET-pIRIR signal sensitivity. *Quat. Geochronol.* **20**, 51–64 (2014).
- Galbraith, R. Graphical display of estimates having differing standard errors. *Technometrics* **30**, 271–281 (1988).
- Galbraith, R. F., Roberts, R. G., Laslett, G. M., Yoshida, H. & Olley, J. M. Optical dating of single and multiple grains of quartz from Jinmium rock shelter, northern Australia: part I, experimental design and statistical models. *Archaeometry* **41**, 339–364 (1999).
- Galbraith, R. & Roberts, R. G. Statistical aspects of equivalent dose and error calculation and display in OSL dating: an overview and some recommendations. *Quat. Geochronol.* **11**, 1–27 (2012).
- Li, B., Roberts, R. G., Jacobs, Z., Li, S.-H. & Guo, Y.-J. Construction of a 'global standardised growth curve' (gSGC) for infrared stimulated luminescence dating of K-feldspar. *Quat. Geochronol.* **27**, 119–130 (2015).
- Li, B., Roberts, R. G. & Jacobs, Z. On the dose dependency of the bleachable and non-bleachable components of IRSL from K-feldspar: improved procedures for luminescence dating of Quaternary sediments. *Quat. Geochronol.* **17**, 1–13 (2013).
- Auclair, M., Lamothe, M. & Huot, S. Measurement of anomalous fading for feldspar IRSL using SAR. *Radiat. Meas.* **37**, 487–492 (2003).
- Huntley, D. J. & Lian, O. B. Some observations on tunnelling of trapped electrons in feldspar and their implications for optical dating. *Quat. Sci. Rev.* **25**, 2503–2512 (2006).

55. Kars, R. H., Wallinga, J. & Cohen, K. M. A new approach towards anomalous fading correction for feldspar IRSL dating—tests on samples in field saturation. *Radiat. Meas.* **43**, 786–790 (2008).
56. Li, B. & Li, S.-H. Investigations of the dose-dependent anomalous fading rate of feldspar from sediments. *J. Phys. D* **41**, 225502 (2008).
57. Kirschvink, J. L. The least-squares line and plane and the analysis of paleomagnetic data. *Geophys. J. R. Astron. Soc.* **62**, 699–718 (1980).
58. Lurcock, P. C. & Wilson, G. S. PuffinPlot: a versatile, user-friendly program for paleomagnetic analysis. *Geochem. Geophys. Geosyst.* **13**, Q06Z45 (2012).
59. Torsvik, T., Briden, J. & Smethurst, M. Super-IAPD interactive analysis of palaeomagnetic data. <http://www.geodynamics.no/resources.html> (2000).
60. Butler, R. F. *Paleomagnetism: Magnetic Domains to Geologic Terranes* Ch. 6, 103–120 (Blackwell Scientific, 1992).
61. Rivera, T. A., Storey, M., Schmitz, M. D. & Crowley, J. L. Age intercalibration of $^{40}\text{Ar}/^{39}\text{Ar}$ sanidine and chemically distinct U/Pb zircon populations from the Alder Creek Rhyolite Quaternary geochronology standard. *Chem. Geol.* **345**, 87–98 (2013).
62. Storey, M., Roberts, R. G. & Saidin, M. Astronomically calibrated $^{40}\text{Ar}/^{39}\text{Ar}$ age for the Toba supereruption and global synchronization of late Quaternary records. *Proc. Natl Acad. Sci. USA* **109**, 18684–18688 (2012).
63. Elburg, M. A. & Foden, J. Geochemical response to varying tectonic settings: an example from southern Sulawesi (Indonesia). *Geochim. Cosmochim. Acta* **63**, 1155–1172 (1999).
64. Wilson, M. E. J., Bosence, D. W. J. & Limbong, A. Tertiary syntectonic carbonate platform development in Indonesia. *Sedimentology* **47**, 395–419 (2000).



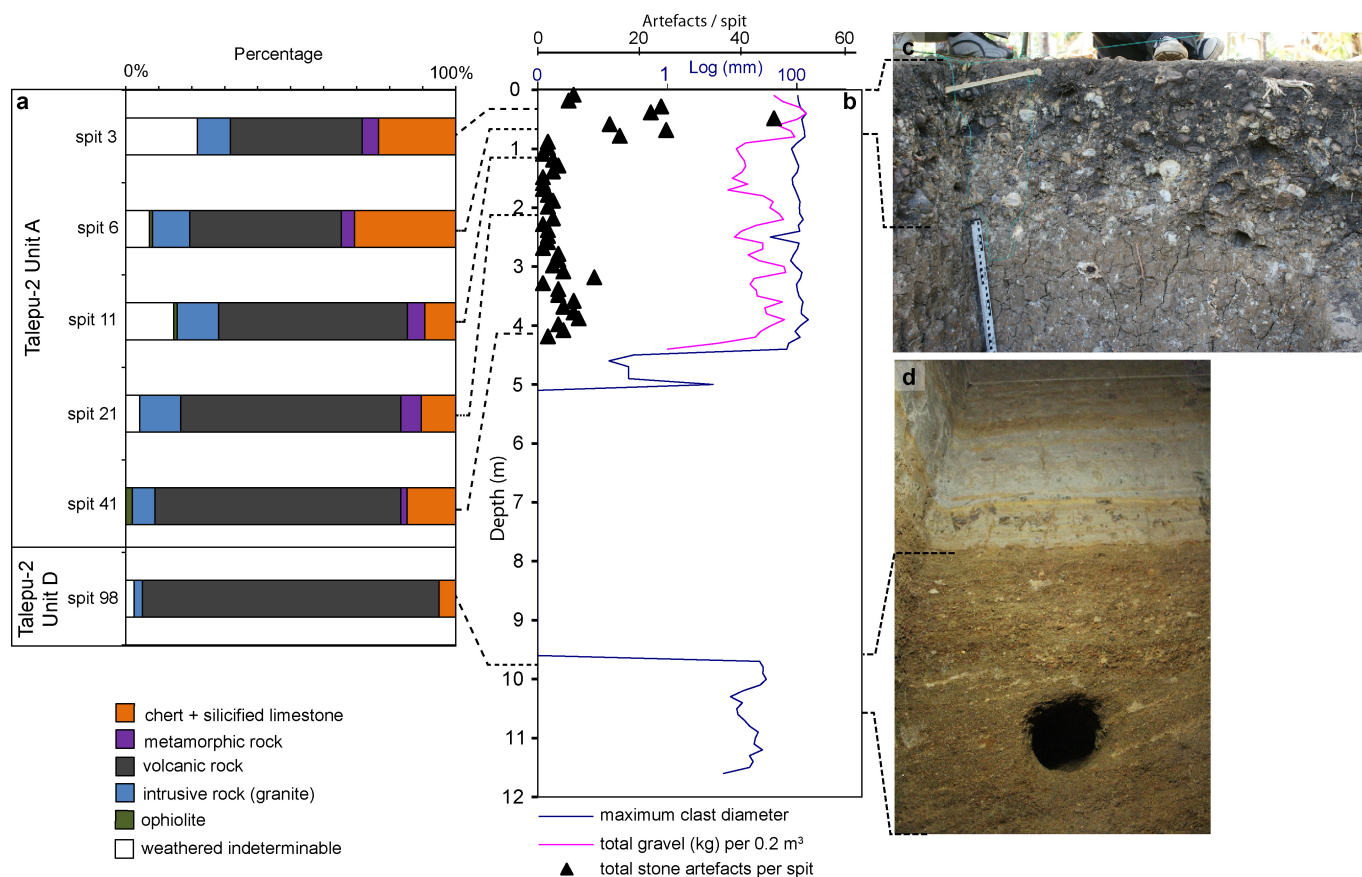
Extended Data Figure 1 | Research location and geology of southwest Sulawesi and Talepu. **a**, Digital elevation model of southwest Sulawesi. 1: Talepu; 2: Maros karst area. Area enclosed by rectangle shown in **d**; map data: copyright USGS/NASA SRTM (2007). **b**, Geological map of southwest Sulawesi (data from refs 11, 63, 64). The Walanae Depression (WD) is an elongate fault-bounded basin (also known as the West Sengkang Basin) separated from the Bone Mountains to the east by a major fault, the East Walanae fault (EWF), which formed in response to east-west compression and strike-slip movements along the Walanae fault zone. To the west the basin is bordered by the Western Dividing Range, consisting of uplifted Miocene volcanics deposited in a shallow marine environment¹³. The Walanae Depression basin infill consists of a several-kilometre-thick regressive sequence, named the Walanae Formation. In the southern part of the Walanae Depression, the Walanae Formation is folded and deformed by Pleistocene compression, whereas to the north near Lake Tempe deposition continues to the present day. 1: Talepu; 2: Paroto (alluvial terrace of the Walanae River); 3: Beru⁸; 4: Tanrung River (palaeontological site: coastal terrace deposits¹¹); 5: Sompe⁸⁻¹⁰; 6: Celeko⁸⁻¹⁰; 7: Maros karst archaeological sites⁵. **c**, Schematic stratigraphic scheme for the northern Sengkang Basin at the latitude of the Talepu site (green dotted line in **b**). The Walanae Formation basin fill represents a regressive sequence that was strongly influenced by tectonic movements along the Walanae fault zone. The youngest unit of the Walanae Formation is the Beru member (deltaic sands, clays and gravels), which contains fossil vertebrate remains of the Walanae Fauna¹¹.

The lower part of the Beru member (unit A) is characterized by sedimentary structures indicative of shallow marine/estuarine/fluvial depositional environments. The upper part of the Beru member (unit B) consists of fully terrestrial fluvio-lacustrine deposits, which merge into the modern floodplain along the depocentral axis of the Walanae Depression. The coarser-grained unit B of the Beru member was not deposited in the Sengkang anticline, which started to rise during the Middle Pleistocene, or in the southern portion of the Walanae Depression south of Talepu. East of the Walanae fault zone, in the East Sengkang Basin, uplift and folding during the Pliocene caused a depositional hiatus. Here Late Miocene deformed marine deposits of the Walanae Formation are unconformably overlaid by a conglomerate up to 5 m thick, the Tanrung Formation, which contains a distinct fossil vertebrate fauna¹¹. During the Middle and Late Pleistocene, uplift of the Western Dividing Range generated the formation of alluvial fans and influxes of coarse-grained boulder conglomerates into the Walanae Depression. **d**, Geological map of Talepu area with sub-horizontal layering, fault-bounded to the east by steeply west-dipping strata of the Sengkang anticline. 1: Modern alluvium; 2: Late Pleistocene alluvial terrace; 3-7: lithological sub-units of the Walanae Formation: 3: fluvio-lacustrine facies of the upper part of the Beru member; 4: fluvio-estuarine facies of the lower part of the Beru member; 5: shallow marine facies of the Samaoling member; 6: deep marine facies of the Burecing member; 7: coral reef facies of the Tacipi member; 8: strike and dip; 9: sub-horizontal layering; 10: major fault; 11: sites with surface-collected stone artefacts; 12: sites with *in situ* stone artefacts; 13: fossil vertebrate localities.



Extended Data Figure 2 | Talepu site and excavation images. **a**, View to the east of the north–south trending ridge. Talepu is located behind the palm trees on the left. **b**, View towards the East Baulk of Talepu excavation 2 in 2009. In 2010 the 1 m × 2 m excavation was extended to a 2 m × 2 m excavation. **c**, The 12 m deep Talepu excavation 2 in 2012. **d**, Talepu

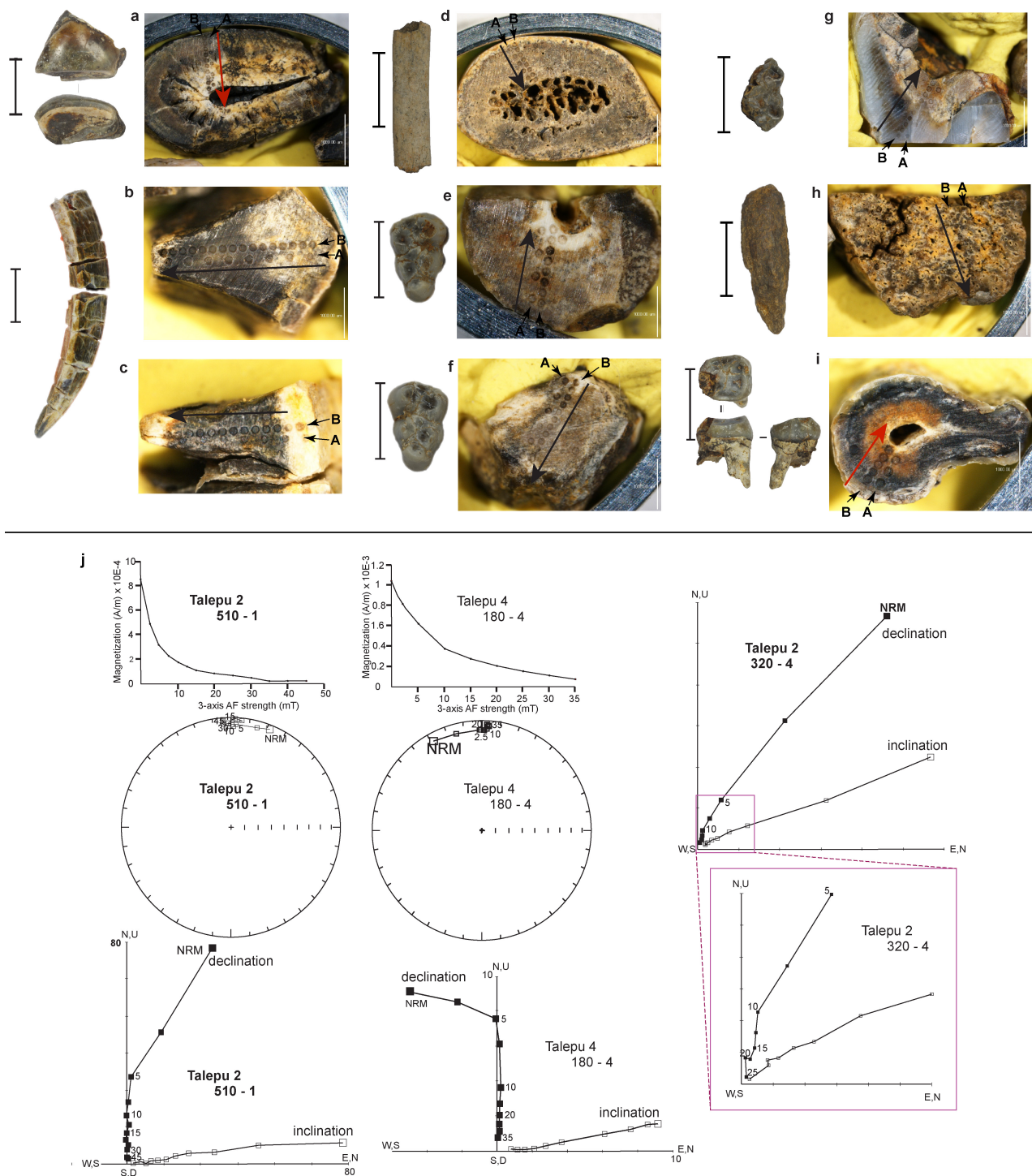
excavation 4 in 2010. View towards the North Baulk. **e**, Talepu excavation 2 in 2012. Photograph shows the North Baulk at 4–5 m depth, with, in the upper part, the base of gravel unit A₇ and the three holes left by sampling for optical dating.



Extended Data Figure 3 | Composition of gravels and distribution of stone artefacts in T2, and size distribution of recent anoa molars.

a, Gravel compositions based on pebble counts (200 pebbles per level). Overall, the composition of the gravel is dominated by volcanic pebbles, which become more abundant with increasing depth, probably as a result of less intense weathering further down (near the surface, the volcanic clasts are frequently weathered to a crumbly clayey 'ghost'). Note the increase both in weathering-resistant silicified rock pebbles and in heavily weathered indeterminate clasts towards the top of the sequence. **b**, Total number of artefacts per 10 cm spit (black triangles), total amount of gravel clasts per spit (red graph) and the maximum clast diameter per spit (blue graph: values represent the mean maximum clast diameter of the ten largest clasts). Note the higher concentration of both gravel and artefacts in the topsoil: pebbles and artefacts are concentrated by winnowing of

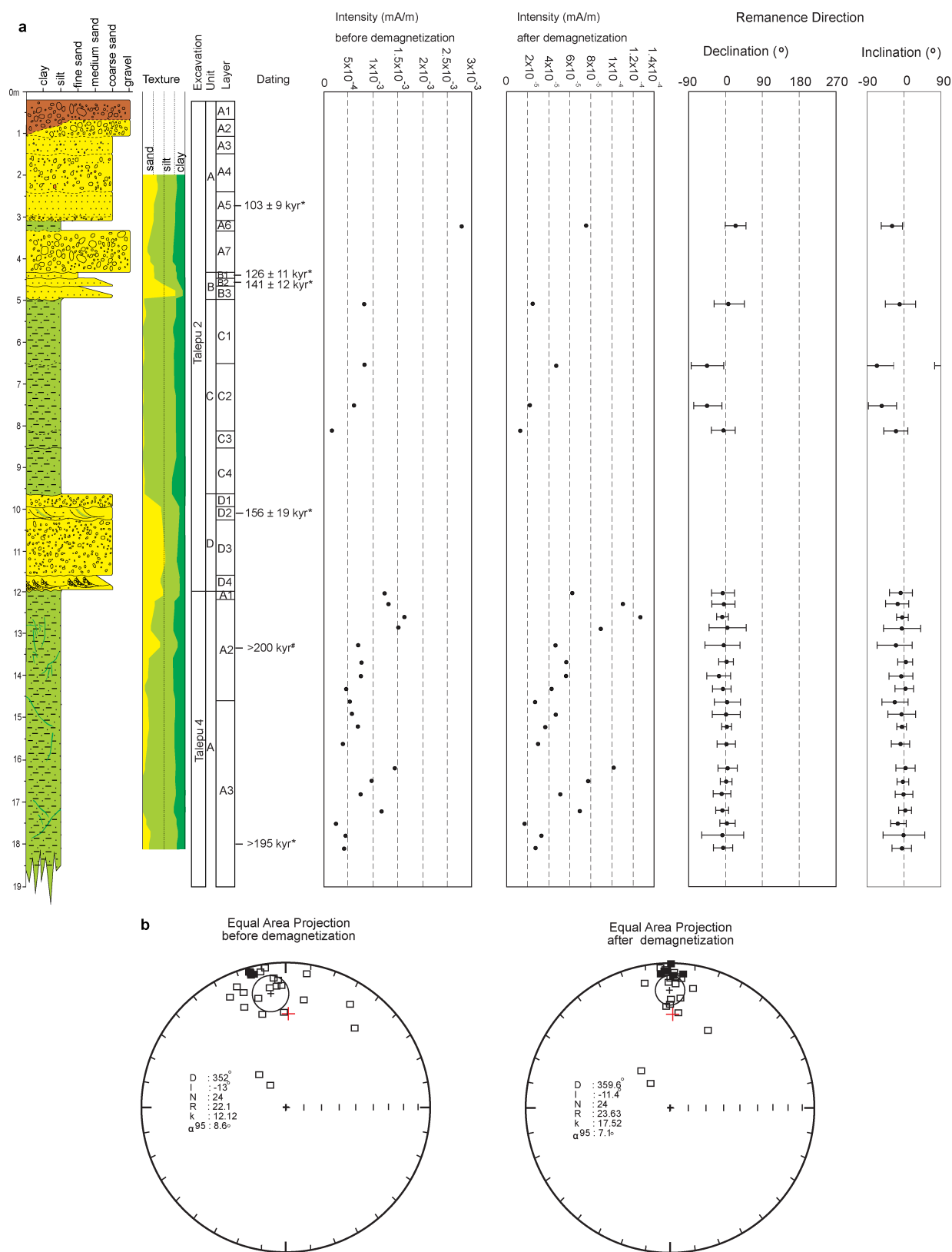
sand and clay by sheetwash processes. **c**, Detail of the topsoil as exposed in the west baulk of excavation T2. **d**, Detail of the basal sequence as exposed in the north baulk of excavation T2. Note the cross-bedded foresets of the pebbly sand of sub-unit D₂, with inter-bedded mud laminae, indicative of tidal activity. Diameter of the round sample hole (for optical dating) is ~10 cm. **e**, Histogram of the transverse diameter measurements (in millimetres) of recent lowland anoa (*Bubalus [Anoa] depressicornis*) lower molars measured in the collections of the Naturalis Biodiversity Centre, Leiden, the Netherlands ($n = 32$). The lowland anoa is the largest living anoa, bigger in body size than the mountain anoa, *Bubalus [Anoa] quarlesi*. The lower molar fragment from unit A of the Talepu-2 excavation (Fig. 3t) has a preserved basal transverse diameter of 14.4 mm and an estimated basal transverse diameter of 15.5 mm, slightly above the size range for extant lowland anoa.



Extended Data Figure 4 | Fossil samples used for uranium-series dating and demagnetization results of representative palaeomagnetic samples.

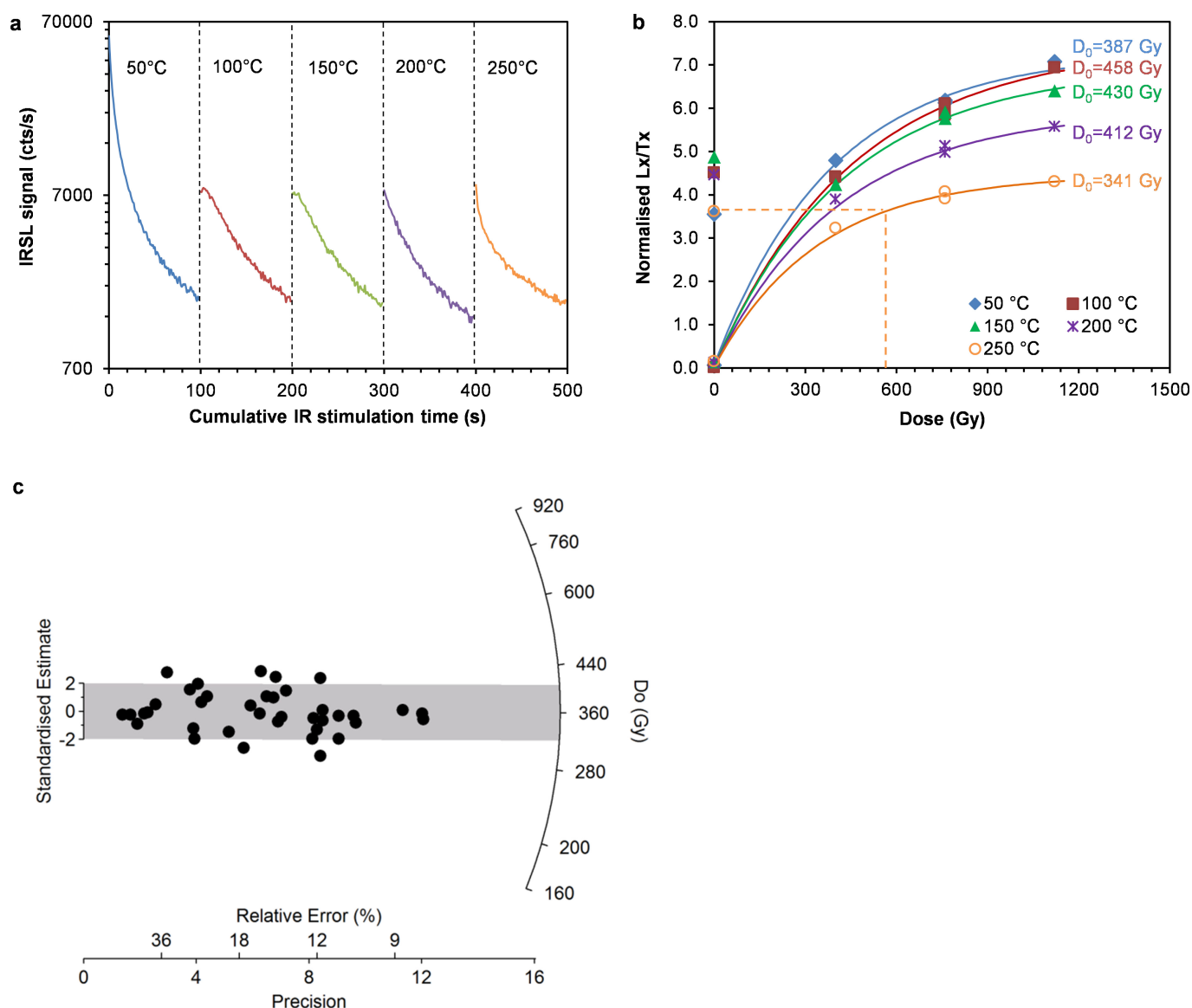
Faunal remains from Talepu excavation 4, used for uranium-series laser ablation dating (a–i) and representative NRM intensity plots of progressive demagnetization (j). a–i, Close-ups of the surface sections for each fossil with the laser spot profiles. All fossils originate from excavation T4, sub-unit E2. Scale bars next to fossils are 2 cm; white scale bars in close-ups are 1 mm. Numbers between brackets are Australian National University laboratory numbers. a, Specimen TLP10-F8 (ANU-2946), *Celebochoerus* upper left first incisor; laser ablation transect on cut section of root. b, c, Two sections measured on different transects of the same specimen, TLP10-F1 (ANU-2947 and ANU-2948), a *Celebochoerus* lower left canine. d, Specimen TLP10-F7 (ANU-2951), rib fragment of *Celebochoerus*; e, Specimen TLP10-F4 (ANU-2954), *Celebochoerus* upper left third molar; laser ablation transect on cut section of root;

f, Specimen TLP10-F3 (ANU-2956), *Celebochoerus* upper right third molar (same individual as previous); laser ablation transect on cut section of root. g, Specimen TLP10-F9 (ANU-2955), *Celebochoerus* upper molar fragment; laser ablation transect on cut section of enamel and dentine. h, Specimen TLP10-F6 (ANU-2949), bone fragment. i, TLP10-F2 (ANU-2942), *Celebochoerus* upper left fourth premolar; laser ablation transect on cut section of root. j, NRM intensity plot of progressive demagnetization (upper left), equal area projections (middle left) and vector end-point demagnetization orthogonal plots (bottom left) for two Talepu palaeomagnetic samples (T2-510-1, T4-180-4). To the right the demagnetization curve for an additional sample (T2-320-4) is given. The inset shows the zoomed-out trajectory endpoints of sample T2-320-4. Open squares on the equal area projection diagrams indicate an upper hemisphere magnetic direction.



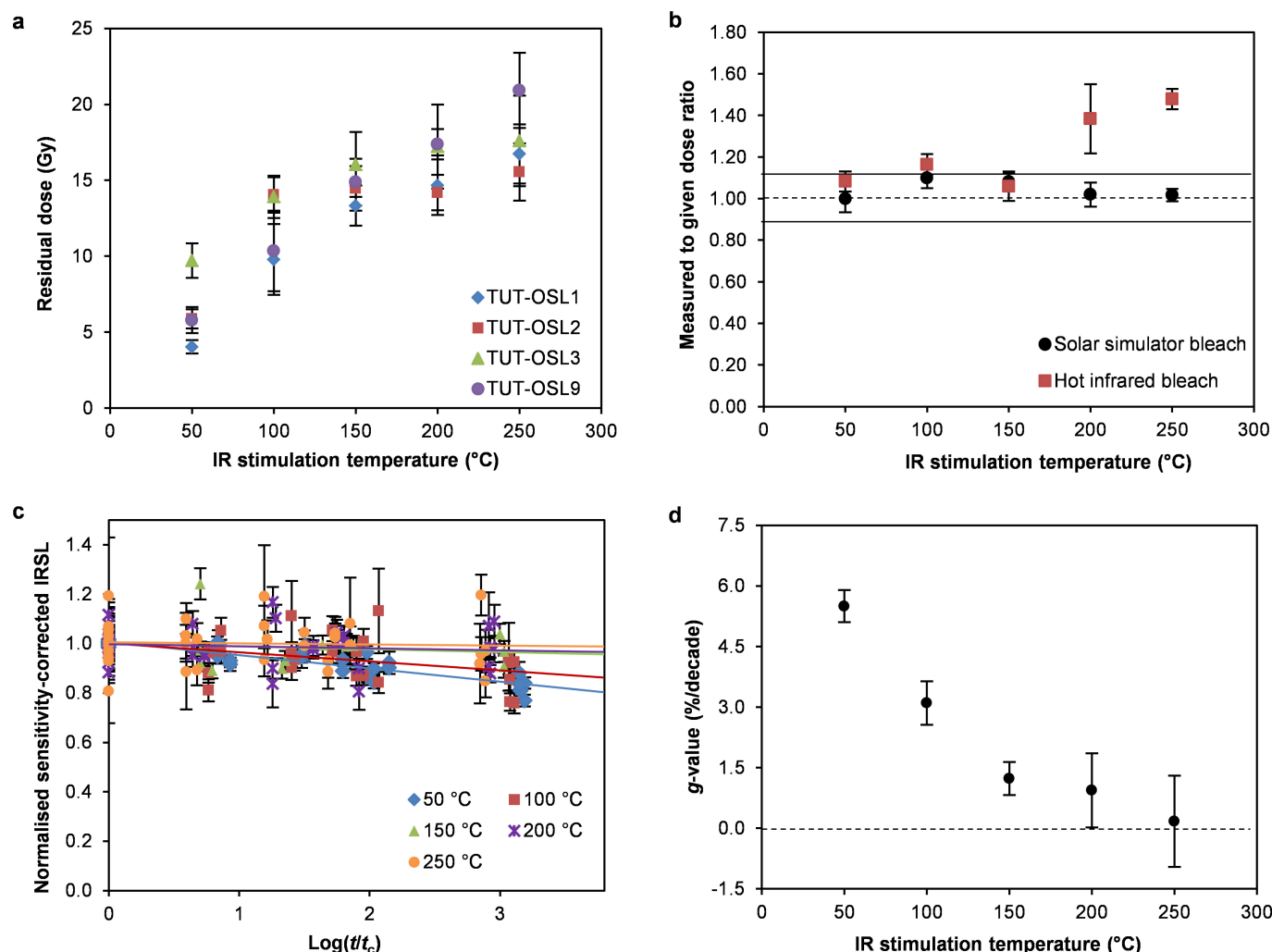
Extended Data Figure 5 | Lithological and magnetic properties against depth for the composite stratigraphic column at Talepu. a, Columns from left to right show lithology, sand/silt/clay ratios, NRM magnetic intensities before and after demagnetization, and magnetic declination and inclination directions. The intensities before and after demagnetization represent averages for each sampled level. Declination and inclination

values are the averages of the higher coercivity stable magnetization (ChRM) with their 95% confidence ranges. **b,** Equal-area projections of NRM and ChRM directions for all sampled levels, and the mean direction (circles with crosses; the mean is of all sampled levels except the two levels with deviating inclinations: $n = 22$) and present-day magnetic direction in the area (red crosses).



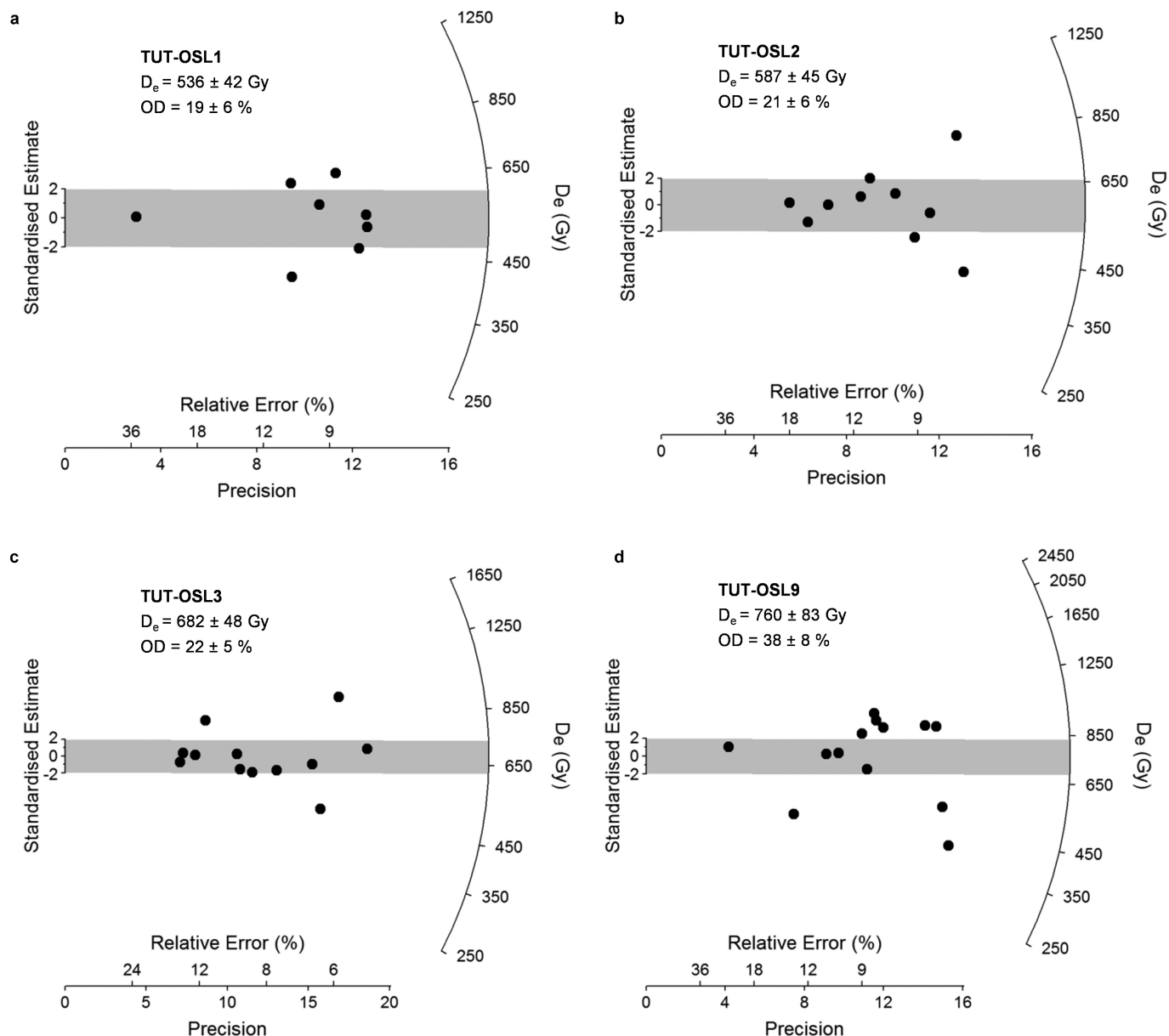
Extended Data Figure 6 | Decay curves and dose-response curves of the IRLS signals for K-feldspars from Talepu. **a**, Representative IRSL and MET-pIRIR decay curves for a single aliquot of sample TUT-OSL2, stimulated at different temperatures (shown above each curve). **b**, Dose-response curves for the IRSL (50°C) and MET-pIRIR (100–250°C) signals for the same aliquot. The natural signals are shown on the x axis using the

same symbols as the regenerative signals. The data points were fitted using a single saturating-exponential function. The best-fit curves are shown as solid lines and the characteristic saturation dose (D_0) values are indicated. **c**, Radial plot showing the D_0 values from 38 aliquots for different samples. The grey band shows the mean of the D_0 values.



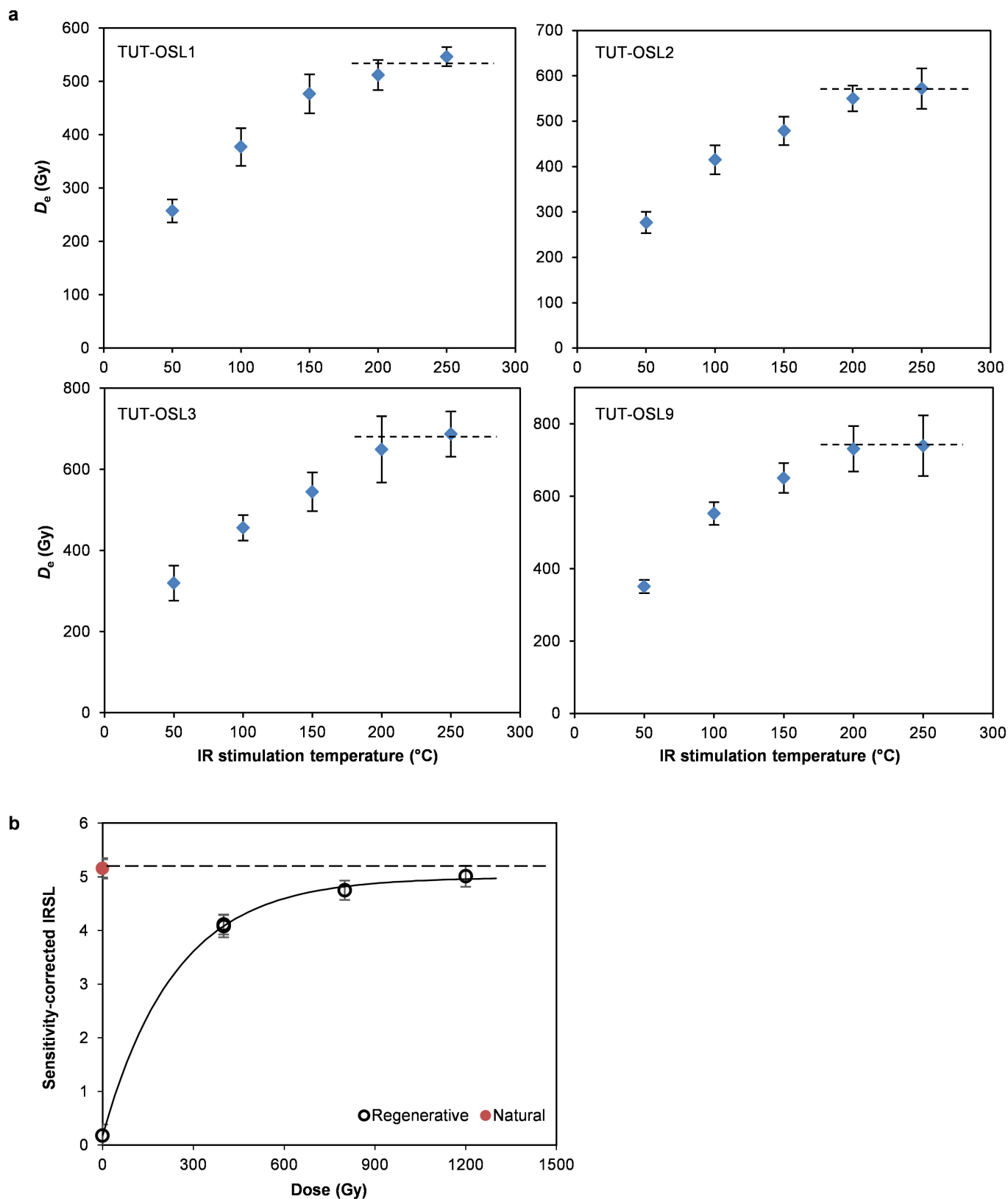
Extended Data Figure 7 | Results from residual dose, dose recovery and anomalous fading tests. **a**, Residual doses measured for bleached aliquots of the four samples from the upper trench, plotted against stimulation temperature. Each data point represents the mean and standard error for four aliquots. **b**, Results of the dose recovery test conducted on sample TUT-OSL1. The measured/given dose ratios are shown for the IRSL and MET-pIRIR signals at the different stimulation temperatures. Each data point represents the mean and standard error for four aliquots. The data shown in red squares were obtained using a hot IR bleach at the end of each SAR cycle, as per the conventional MET-pIRIR procedure. The data shown in black circles were obtained with the modified MET-pIRIR procedure (Supplementary Table 4), using a solar simulator bleach instead of a hot IR bleach. The dashed line denotes a ratio of unity, and the solid

lines indicate ratios 10% larger and smaller than unity. The data (circles) obtained using the modified MET-pIRIR procedure fall within the latter band. **c**, Decay of the sensitivity-corrected IRSL and MET-pIRIR signals with their standard errors of six aliquots from TUT-OSL3, plotted against $\log(t/t_c)$ where t is the delayed period for each measurement and t_c is the time for the first measurement ($t_c = 720, 870, 1,040, 1,240$ and $1,480$ s for the signals measured at 50, 100, 150, 200 and 250 °C, respectively). The sensitivity-corrected signals were normalized to the first measurements. **d**, Anomalous fading rates (g values) and their standard errors for the IRSL and MET-p IRIR signals of TUT-OSL3 obtained using the data sets in **c**, plotted against stimulation temperature. All the g values have been normalized to a delay time of 2 days.



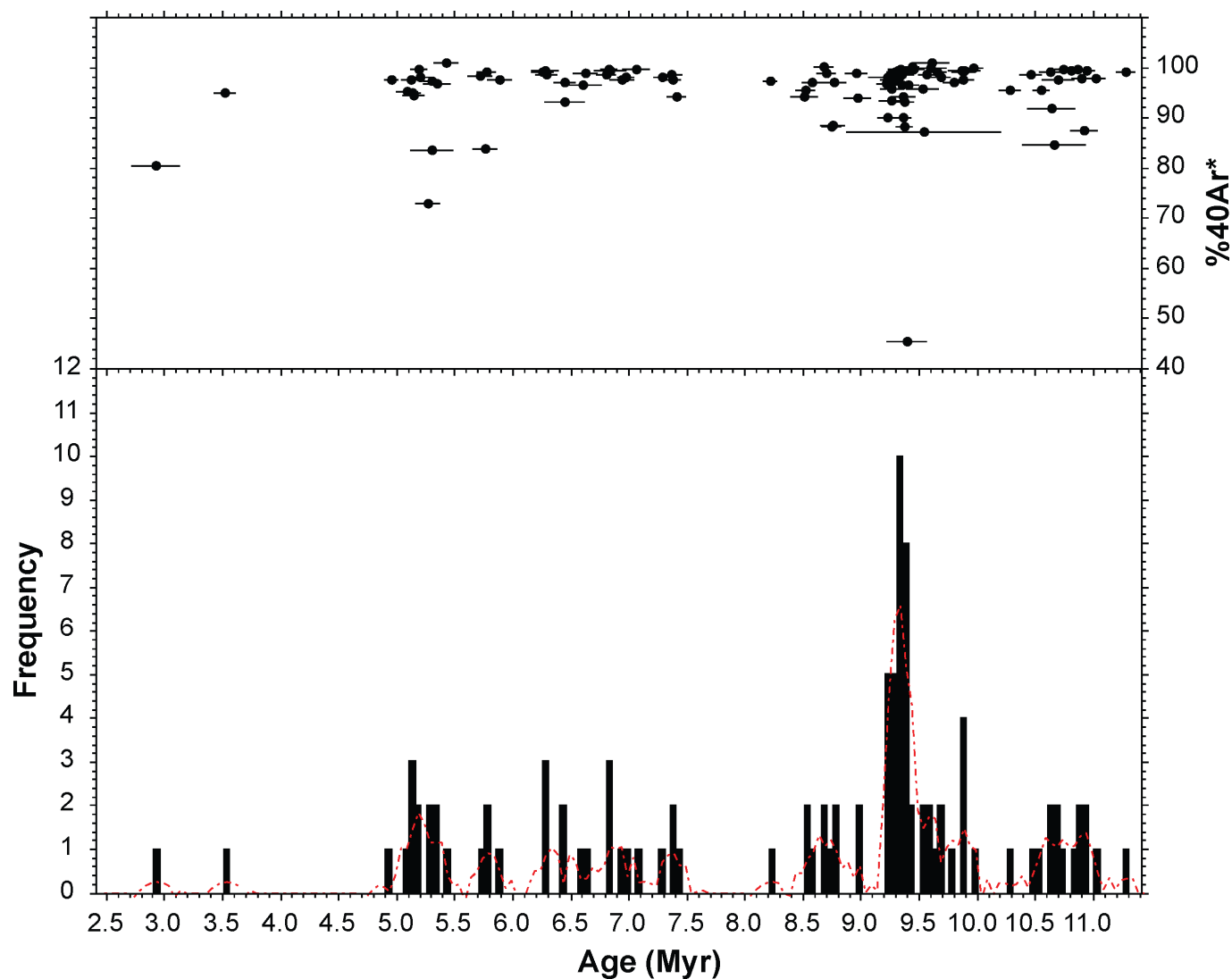
Extended Data Figure 8 | Radial plots of single-aliquot D_e values for the TUT samples. a, TUT-OSL1. b, TUT-OSL2. c, TUT-OSL3. d, TUT-OSL9. The grey band in each plot shows the weighted mean of the measured

D_e values estimated using the central age model. The D_e estimate and the overdispersion (OD) value for each D_e distribution based on the central age model are also shown in each plot.



Extended Data Figure 9 | D_e versus temperature plots for the TUT samples and the dose-response curve for sample TLT-OSL6. a, Plots of the weighted mean D_e against stimulation temperature for the TUT samples. The dashed line in each plot shows the plateau range of D_e values. Each data point represents the mean and standard error for 8 (TUT-OSL1), 10 (TUT-OSL2), 12 (TUT-OSL3) and 13 (TUT-OSL9) aliquots. **b,** Dose-response

curve for the sensitivity-corrected MET-pIRIR 250°C signal from an aliquot of sample TLT-OSL6. The regenerative-dose data points and their standard errors were fitted using a single saturating-exponential function, and the best-fit curve is shown as a full line. The natural signal of this aliquot (red circle on the y axis) falls in the saturated region of the curve (see dashed line), so only a minimum D_e can be estimated.



Extended Data Figure 10 | $^{40}\text{Ar}/^{39}\text{Ar}$ fusion ages of single sanidine crystals from sample TAL-10-01 (T4, sub-unit E₂, 2.5 m below surface). Top: the individual ages $\pm 1\sigma$. Bottom: histogram based on over 100 analyses. Excluding the two youngest ages, the crystal ages span a period

from 5 to 11 million years ago, but with the main age population at about 9.4 million years ago. The ages correspond with the Late Miocene collision phase, when potassium-rich volcanics of the Camba Formation were formed along the Western Dividing Range⁶³.

Diet-induced extinctions in the gut microbiota compound over generations

Erica D. Sonnenburg^{1*}, Samuel A. Smits^{1*}, Mikhail Tikhonov^{2,3}, Steven K. Higginbottom¹, Ned S. Wingreen^{4,5} & Justin L. Sonnenburg¹

The gut is home to trillions of microorganisms that have fundamental roles in many aspects of human biology, including immune function and metabolism^{1,2}. The reduced diversity of the gut microbiota in Western populations compared to that in populations living traditional lifestyles presents the question of which factors have driven microbiota change during modernization. Microbiota-accessible carbohydrates (MACs) found in dietary fibre have a crucial involvement in shaping this microbial ecosystem, and are notably reduced in the Western diet (high in fat and simple carbohydrates, low in fibre) compared with a more traditional diet³. Here we show that changes in the microbiota of mice consuming a low-MAC diet and harbouring a human microbiota are largely reversible within a single generation. However, over several generations, a low-MAC diet results in a progressive loss of diversity, which is not recoverable after the reintroduction of dietary MACs. To restore the microbiota to its original state requires the administration of missing taxa in combination with dietary MAC consumption. Our data illustrate that taxa driven to low abundance when dietary MACs are scarce are inefficiently transferred to the next generation, and are at increased risk of becoming extinct within an isolated population. As more diseases are linked to the Western microbiota and the microbiota is targeted therapeutically, microbiota reprogramming may need to involve strategies that incorporate dietary MACs as well as taxa not currently present in the Western gut.

The gut microbiota of hunter-gatherers and populations consuming a rural agrarian diet is distinct, and contains greater diversity than the microbiota of Westerners^{4–9} (Extended Data Fig. 1). One possible explanation for the greater microbiota diversity seen in hunter-gatherers and agrarians is the large quantity of dietary fibre they consume relative to Westerners^{4,6,10,11}. MACs, which are abundant in dietary fibre, serve as the primary source of carbon and energy for the distal gut microbiota^{3,12}. Therefore, we sought to determine whether a diet low in MACs could drive loss of taxa within the gut microbiota.

Humanized mice (4 weeks old, $n = 10$) were fed a diet rich in fibre derived from a variety of plants (high-MAC) for 6 weeks, and randomly divided into two groups (Extended Data Fig. 2). One group was switched to a low-MAC diet for 7 weeks, after which they were returned to the high-MAC diet for 6 weeks (Fig. 1a and Extended Data Table 1). The control group was maintained on the high-MAC diet throughout the experiment. At the start of the experiment, the microbiota composition from both groups of mice was indistinguishable ($P = 0.2$, Student's t -test; UniFrac distance; no significant difference in operational taxonomic unit (OTU) frequency observed between groups, Mann-Whitney U test). The diet-switching mice, while consuming the low-MAC diet, had an altered composition relative to controls ($P = 10^{-25}$, Student's t -test; UniFrac distance). Weeks after returning to the high-MAC diet, the microbiota of the diet-switching mice remained

distinct from controls ($P = 3 \times 10^{-8}$, Student's t -test; UniFrac distance at 15 weeks) (Fig. 1b). To determine whether taxa had been lost over the course of the diet perturbation, we focused on a subset of OTUs that met stringent measures of prevalence and abundance and could be confidently monitored over the course of the experiment ('high-confidence' OTUs, see Methods). We identified 208 high-confidence OTUs in the diet-switching group and 213 high-confidence OTUs in the control group (Extended Data Table 2). When mice were switched from the high-MAC diet to the low-MAC diet, we observed that 60% of taxa (124 out of 208) decreased in abundance at least fourfold compared with only 11% of the control group (25 out of 213) (Supplementary Table 1). When these mice were returned to a high-MAC diet, 33% (71 out of 208) were fourfold less abundant. The control group did not change significantly (10% were fourfold less abundant; 22 out of 213) (Fig. 1c and Supplementary Table 2). These data reveal two divergent qualities of the microbiota. First, 59 of the 208 high-confidence OTUs that exhibit diet-induced decline in abundance recovered (were no longer at least twofold less abundant) with the reintroduction of MACs illustrating microbiota resilience over short time scales (Supplementary Table 1). Second, however, the low-MAC-diet perturbation induced 'scars' on the microbiota.

We proposed that diet-induced microbiota diversity loss would be magnified over generations. Mice from the previous experiment consuming the low-MAC diet or the high-MAC diet were used to generate a litter of pups. Pups were weaned onto the respective diets of their parents. This breeding strategy was repeated for four generations. For each generation, low-MAC-diet parents were switched to the high-MAC diet after their pups were weaned, to see whether taxa that became undetectable while MACs were scarce would bloom in the presence of MACs (Fig. 2a and Extended Data Fig. 3). At 5 weeks old, mice propagated in the low-MAC diet condition (born to low-MAC-diet parents and consuming a low-MAC diet) had a lower-diversity microbiota than high-MAC-fed controls ($P = 3 \times 10^{-6}$, $P = 8 \times 10^{-5}$ and $P = 8 \times 10^{-6}$, Student's t -test, Shannon index; generations two, three and four, respectively) (Fig. 2b, top). Even after mice were switched to the high-MAC diet for several weeks, their microbiota diversity did not recover to control levels ($P = 2 \times 10^{-6}$, $P = 1 \times 10^{-8}$ and $P = 1 \times 10^{-4}$, Student's t -test; generations two, three and four, respectively, at week 15) (Fig. 2b, bottom). With each generation, the microbiota composition of the diet-switching group showed increasing departure from that of controls (Fig. 2c). Weaning the diet-switching lineage directly onto the high-MAC diet did not correct the diversity loss relative to controls ($P = 3 \times 10^{-6}$ Student's t -test, Shannon index), and there was no difference in composition between this group and the generation-four diet-switching group ($P = 0.9$, Student's t -test; UniFrac distance; week 13, generation four, and week 5, generation five) (Extended Data Fig. 4).

Plotting the relative abundance of the high-confidence OTUs over time revealed a pattern of taxa loss over generations in the

¹Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305, USA. ²Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. ³Kavli Institute for Bionano Science and Technology, Harvard University, Cambridge, Massachusetts 02138, USA. ⁴Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, USA. ⁵Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA.

*These authors contributed equally to this work.

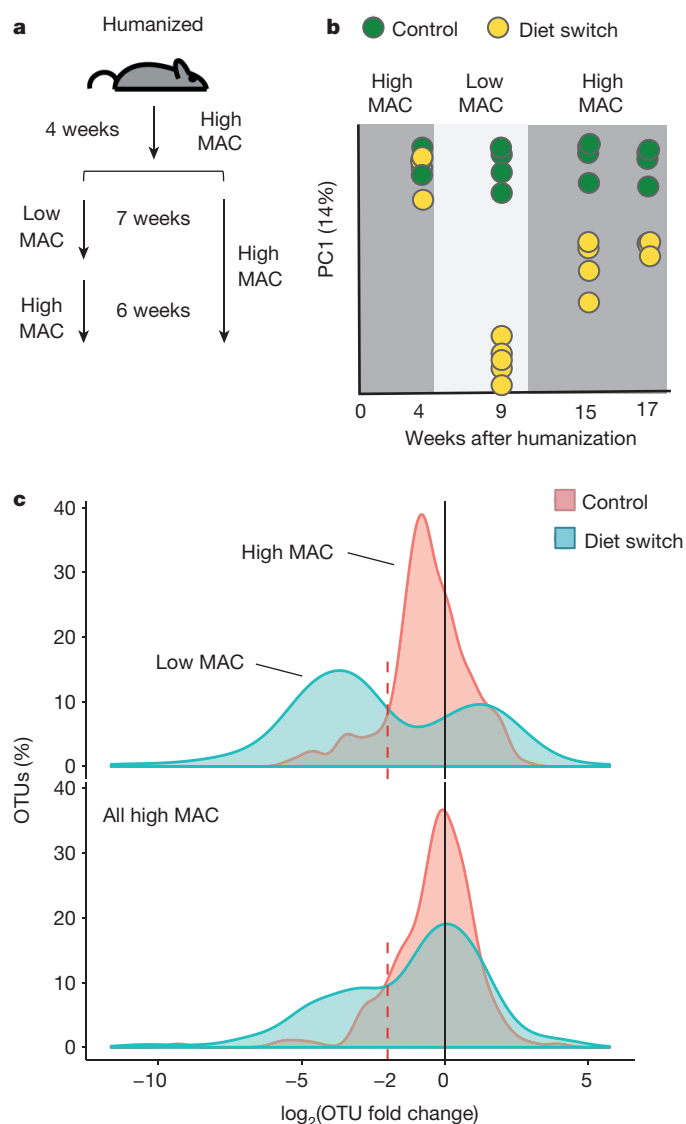


Figure 1 | Taxa reduction observed in low-MAC diet is largely reversible in a single generation. **a**, Schematic of mouse experiment. Humanized mice ($n = 10$) were maintained on a high-MAC diet for 4 weeks, after which one-half of the mice were switched to a low-MAC diet for 7 weeks. These mice were then switched back to the high-MAC diet for 6 weeks. **b**, Principle coordinate (PC) analysis of the UniFrac distance for 16S ribosomal RNA amplicon profiles from faecal samples collected from the diet-switching mice (yellow, $n = 5$) and control high-MAC-diet mice (green, $n = 5$). **c**, Distribution of OTU fold changes for diet-switching (blue, $n = 5$) or control (red, $n = 5$) groups comparing baseline (4 weeks post-humanization) versus week 9 (5 weeks post-low-MAC diet for 'diet switch' group; top panel) and baseline versus week 15 (4 weeks after return to high-MAC diet for 'diet switch' group, bottom panel).

diet-switching group (Fig. 2d). Specifically, in the diet-switching group, generations one, two and three exhibited a progressive loss in high-confidence OTUs while consuming the low-MAC diet (72%; 150 out of 208 lost by generation four, week 15) (Supplementary Table 3). In each generation, switching to the high-MAC diet allowed for the recovery of a small number of taxa (grey versus yellow highlighted rows within each generation in Fig. 2d), but most did not return (141 out of 208 were undetectable by generation four, week 15) (Supplementary Table 3). Most of the lost taxa (112 out of the 141) were from the Bacteroidales order with an additional loss of 26 taxa from the Clostridiales, making the Clostridiales the most numerous high confidence taxon present in the fourth generation (Extended Data Fig. 5a, b).

To determine whether the carbohydrate degrading capacity of the microbiota had been altered over the four generations, we compared imputed glycoside hydrolases between the first and fourth generations of both the diet-switching and control groups after validating this method^{13,14} (Supplementary Table 4). Although representation of glycoside hydrolase families is not a perfect correlate of specificity for polysaccharide degradation (for example, owing to combinatorial activity or polyspecificity within a family), loss of representation within glycoside hydrolase families provides one measure of changes in glycan-degrading capacity. Twenty-two glycoside hydrolase families showed a loss in abundance in the diet-switching group between the initial time point in generation one versus generation four, 4 weeks after the switch to the high-MAC diet ($P < 0.05$ plus at least a twofold change, Bonferroni-corrected Student's t -test) (Extended Data Fig. 5c and Supplementary Table 5). No differences in glycoside hydrolase families were observed in the control group. An overall loss in glycoside hydrolase diversity occurred between generation-four diet-switching mice on the high-MAC diet relative to generation-one mice ($P = 0.0002$, Shannon index of glycoside hydrolase subfamilies; Supplementary Table 6 and Extended Data Table 3). No difference in glycoside hydrolase subfamilies was observed in the control group. These data demonstrate that, in addition to the loss of high-confidence OTUs, the diet-switching group sustained a widespread and marked loss in glycoside hydrolase repertoire over the four generations. Future experiments will be needed to reveal the functional consequences of these observations in terms of fibre-degrading capacity.

We next wanted to determine whether low-abundance OTUs that bloom when MACs are reintroduced are more likely to be lost owing to inefficient inter-generational transmission. Low-abundance taxa (average abundance < 25 reads) in a given generation were less efficiently transferred to the next generation (average abundance > 10 reads to be considered present, 4 weeks after high-MAC diet) compared to high-abundance taxa (average abundance > 25 reads) ($P = 0.002$, $P = 4 \times 10^{-15}$ and $P = 0.01$, hypergeometric distribution inheritance between generations one and two, two and three, and three and four, respectively) (Supplementary Table 7). Notably, overall diversity and composition did not change between the third and fourth generations (Fig. 2b, c); however, we wondered whether additional loss could be obscured owing to lack of resolution of OTUs. Therefore, we identified 280 high-confidence sub-OTUs in the control group and 261 high-confidence sub-OTUs in the diet-switching group using a cluster-free filtering approach¹⁵ (Supplementary Table 8). A similar decline in the number of sub-OTUs with each generation was observed (114 out of 261 sub-OTUs were undetectable by generation four, week 15) (Extended Data Fig. 6 and Supplementary Table 8). Most of the lost taxa (77 out of the 114) were from the Bacteroidales order with an additional loss of 32 taxa from the Clostridiales. Between generation three and four, we detected loss of 22 sub-OTUs, compared to four using the lower-resolution high-confidence OTUs (Supplementary Table 8).

Because high dietary MACs were insufficient to restore microbiota composition or diversity to control levels, we tested whether reintroduction of lost bacteria was required. Fourth-generation, diet-switching mice were gavaged with faecal samples (faecal microbiota transplant (FMT) group) from fourth-generation high-MAC-diet controls. Because the low-MAC diet does not support full microbiota diversity (Fig. 2d), the fourth-generation FMT recipients were fed a high-MAC diet for 2 weeks (Fig. 3a). Within 10 days, microbiota composition and diversity of the FMT group was indistinguishable from fourth-generation high-MAC-diet controls ($P = 0.4$, Student's t -test; UniFrac distance; $P = 0.4$, Student's t -test, Shannon index) (Fig. 3b, c), and 110 taxa were restored (average abundance ≥ 1 sequencing read; no taxa restored in no FMT controls) (Fig. 3d and Supplementary Table 9). Restored taxa were predominantly from the Bacteroidales (99 taxa), which experienced the greatest loss in high-confidence OTUs (Extended Data Fig. 7a). Similar results were observed using the high-confidence sub-OTUs (Extended Data Fig. 7b and Supplementary Table 10).

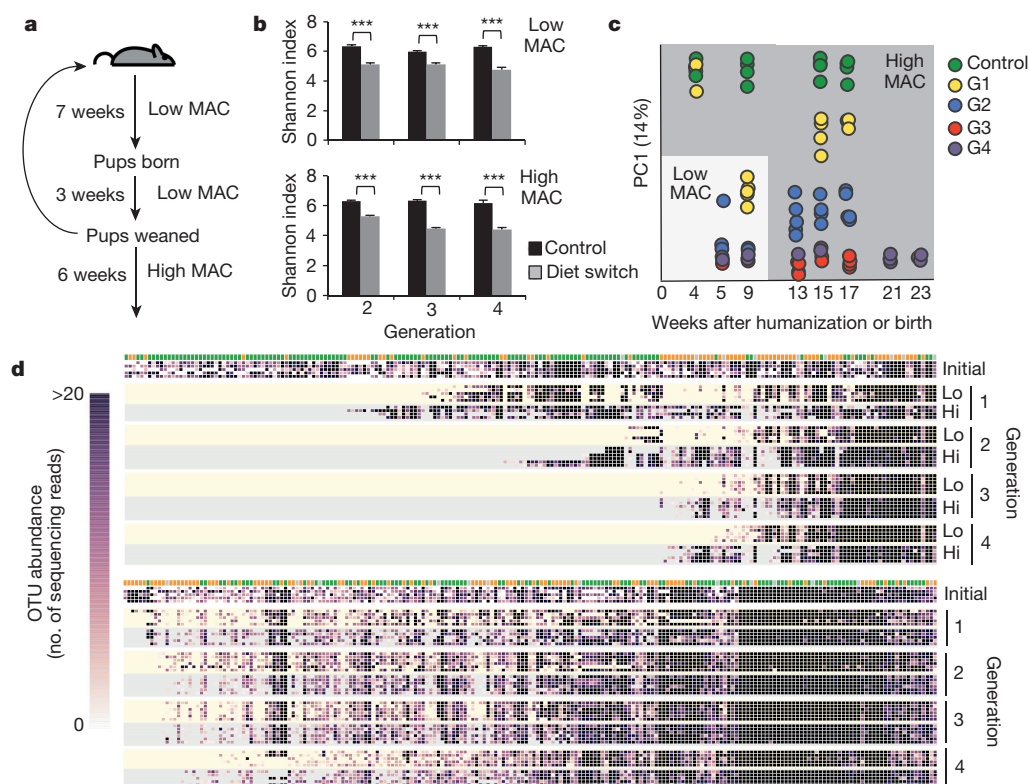


Figure 2 | Inefficient inter-generational transfer of taxa driven to low abundance by low dietary MACs. **a**, Schematic of multigeneration mouse experiment. Second- ($n = 6$), third- ($n = 6$) and fourth- ($n = 6$) generation mice were weaned onto a low-MAC diet. After mice generated a litter of pups that were weaned, low-MAC-diet mice were switched to the high-MAC diet for 6 weeks. A parallel group of control mice were maintained on the high-MAC diet throughout (generation 2, $n = 6$; generation 3, $n = 6$; generation 4, $n = 5$). **b**, Microbiota diversity as measured by Shannon index observed in the microbiota of mice at 5 weeks old (top panel, $n = 6$ for each group) or 4 weeks after shift to high-MAC diet (bottom panel, $n = 6$ for each group) from three generations of diet-switching mice (grey) or control high-MAC-diet mice (black). Error bars are s.e.m. and P values are from two-tailed Student's t -test. **c**, Principal coordinate analysis of UniFrac distance for 16S rRNA amplicon profiles from faecal samples

collected from first-generation mice from the control group consuming a high-MAC diet (green, $n = 5$) or the diet-switching group from generation one (G1; yellow, $n = 5$), two (G2; blue, $n = 6$), three (G3; red, $n = 6$) and four (G4; purple, $n = 6$). **d**, Heat map of abundance of high-confidence OTUs (number of sequencing reads, columns) from the diet-switching group (top) and controls (bottom); taxonomic assignment is indicated at the top of each column (Bacteroidetes, green; Firmicutes, orange; other, grey). Each row represents an individual mouse microbiota from 4 weeks post-humanization (initial), while consuming the low-MAC diet (week 9, lo, shaded yellow), and 4 weeks after switching to the high-MAC diet (week 15, hi, shaded grey). Corresponding time points from controls are similarly shaded. $n = 5$, 6, 6 and 6 for the diet-switching group and $n = 5$, 6, 6 and 5 for the control group for generations one to four, respectively.

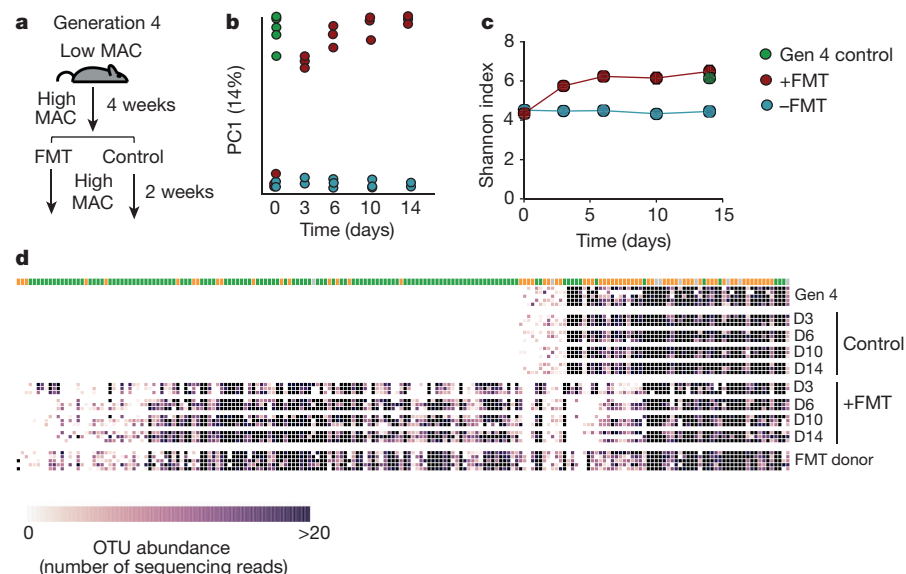


Figure 3 | Reintroduction of lost taxa and a high-MAC diet restores microbiota diversity and composition. **a**, Schematic of faecal transplant mouse experiment. **b**, Principal coordinate analysis of UniFrac distance for 16S rRNA amplicon profiles from faecal samples collected from fourth-generation control mice on a high-MAC diet (green, $n = 6$), fourth-generation diet-switching mice that received a faecal transplant (red, $n = 3$), or did not (blue, $n = 3$). **c**, Microbiota diversity as measured by Shannon index observed in the microbiota of mice that received a faecal transplant (red, $n = 3$) or did not (blue, $n = 3$). A green circle denotes the number of OTUs observed in fourth-generation control mice consuming a high-MAC diet ($n = 6$). Error bars are s.e.m. **d**, Heat map of abundance of high-confidence OTUs (number of sequencing reads) from fourth-generation diet-switching mice ($n = 3$) 3–14 days after FMT, and no-FMT controls ($n = 3$); taxonomic assignment is indicated at the top of each column (Bacteroidetes, green; Firmicutes, orange; other, grey). FMT donor (fourth-generation control mice, $n = 5$) and fourth-generation diet-switching mice ($n = 5$) 4 weeks after consuming high-MAC diet are also shown.

These data demonstrate a diet-induced ratcheting effect in which certain taxa decrease in abundance upon reduced MACs and are not effectively transferred to the next generation. Notably, most of the lost taxa are Bacteroidales, an order that is proficient in consumption of dietary fibre¹⁶. Introduction of dietary MACs are insufficient to regain 'lost' taxa in the absence of their deliberate re-introduction.

Over our history, humans have experienced major dietary changes from gathered to farmed foods during the agricultural revolution, and more recently to the mass consumption of processed foods in the industrialized world. Each dietary shift was probably accompanied by a concomitant adjustment in the microbiota. Here we have used a model in which mice have been colonized with a human microbiota from a Westerner to determine the effect of fibre deprivation over four generations on the gut microbiota. This model does not allow us to address microbiota changes that may have occurred as humans shifted from a hunter-gatherer lifestyle to one from a modern industrialized country. Our data support a model in which consuming a modern diet low in fibre contributes to the loss of taxa over generations, and may be responsible for the lower-diversity microbiota observed in the industrialized world compared to present-day hunter-gatherers and rural agrarians. The data we present also hint that further deterioration of the Western microbiota is possible.

The gut microbiota regulates numerous facets of human biology suggesting that our human genome has been shaped by interactions with these microorganisms over our co-evolutionary history. However, the microbiota can change on a timescale that is much faster than the host allowing for the possibility that the microbiota, if pressed by severe selective forces, could undergo change so rapidly that it cannot be accommodated by our human biology. While the roles of different types of microbiota diversity in host health remain to be defined, it is possible that rewilding the modern microbiota with extinct species may be necessary to restore evolutionarily important functionality to our gut.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 July; accepted 2 December 2015.

1. Hooper, L. V., Littman, D. R. & Macpherson, A. J. Interactions between the microbiota and the immune system. *Science* **336**, 1268–1273 (2012).
2. Karlsson, F., Tremaroli, V., Nielsen, J. & Backhed, F. Assessing the human gut microbiota in metabolic diseases. *Diabetes* **62**, 3341–3349 (2013).
3. Sonnenburg, E. D. & Sonnenburg, J. L. Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab.* **20**, 779–786 (2014).
4. Schnorr, S. L. *et al.* Gut microbiome of the Hadza hunter-gatherers. *Nature Commun.* **5**, 3654 (2014).

5. Yatsunenkov, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222–227 (2012).
6. De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl Acad. Sci. USA* **107**, 14691–14696 (2010).
7. Clemente, J. C. *et al.* The microbiome of uncontacted Amerindians. *Science Advances* **1**, e1500183 (2015).
8. Obregon-Tito, A. J. *et al.* Subsistence strategies in traditional societies distinguish gut microbiomes. *Nature Commun.* **6**, 6505 (2015).
9. Martinez, I. *et al.* The gut microbiota of rural Papua New Guineans: composition, diversity patterns, and ecological processes. *Cell Rep.* **11**, 527–538 (2015).
10. McGill, C. R., Fulgoni, V. L. III & Devereedy, L. Ten-year trends in fiber and whole grain intakes and food sources for the United States population: National Health and Nutrition Examination Survey 2001–2010. *Nutrients* **7**, 1119–1130 (2015).
11. King, D. E., Mainous, A. G. III & Lambourne, C. A. Trends in dietary fiber intake in the United States, 1999–2008. *J. Acad. Nutr. Diet.* **112**, 642–648 (2012).
12. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
13. Kashyap, P. C. *et al.* Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. *Proc. Natl Acad. Sci. USA* **110**, 17059–17064 (2013).
14. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495 (2014).
15. Tikhonov, M., Leach, R. W. & Wingreen, N. S. Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. *ISME J.* **9**, 68–80 (2015).
16. El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D. & Henrissat, B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Rev. Microbiol.* **11**, 497–504 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. St. Onge for technical assistance. This work was funded by a grant from National Institutes of Health NIDDK (R01-DK085025 to J.L.S.), an NSF graduate fellowship (to S.A.S.), a Stanford Graduate Fellowship (to S.A.S.), and the Simons Foundation (to M.T.). J.L.S. holds an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund.

Author Contributions E.D.S. and J.L.S. conceived and designed the project. E.D.S., J.L.S. and S.K.H. designed and supervised the experiments. E.D.S. and S.K.H. performed the experiments. E.D.S., S.A.S. and M.T. analysed the experimental data. N.S.W. designed and supervised data analysis. E.D.S. and S.A.S. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The 16S sequence data have been deposited in the Sequence Read Archive (SRA) under the accession PRJNA303185. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.L.S. (jsonnenburg@stanford.edu).

METHODS

No statistical methods were used to predetermine sample size.

Meta-analysis of human populations. 16S rRNA data sets from Hadza ($n=27$), Malawian and Venezuelan ($n=213$) and American ($n=315$)^{4,5} were trimmed to match FLX chemistry as previously described (MG-RAST Projects 111, 528, 7058)¹². OTUs were picked on the Greengenes 13.8 database with a 97% similarity threshold using UCLUST¹⁷. Alpha- and beta-diversity measures were calculated using unweighted UniFrac¹⁸ on rarefied OTU tables. Principal coordinates were computed using QIIME 1.8 (ref. 19). Faecal 16S rRNA data were not included from other studies of traditional populations because (1) they did not use the V4 region of the 16S rRNA for amplification and were thus not comparable (Papua New Guinea⁹); (2) they have not made the data publicly available (Yanomami⁷); or (3) the control Western data was not comparable to other studies (Matses⁸).

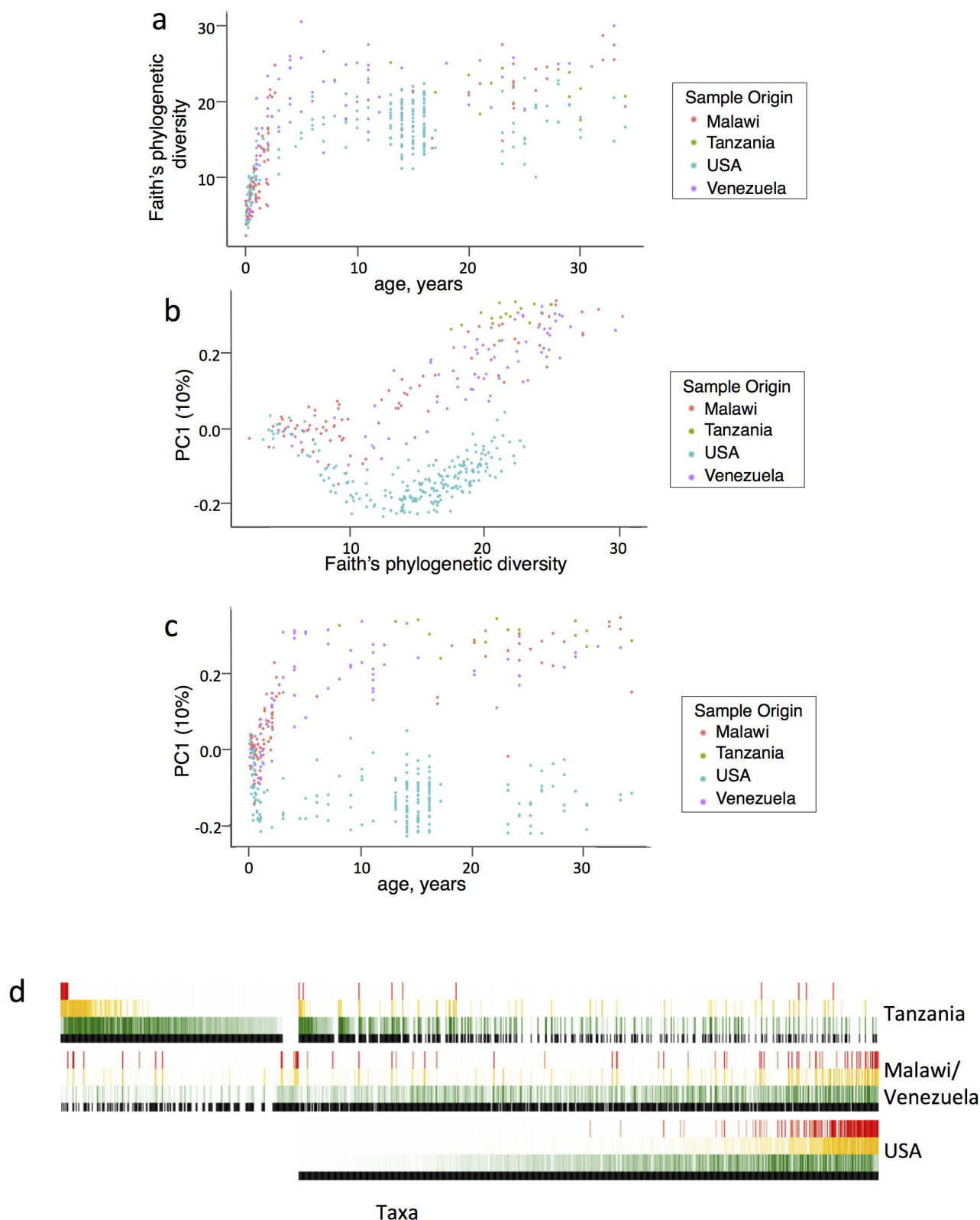
Mice. All mouse experiments were performed in accordance with A-PLAC, the Stanford IACUC on mixed gender germ-free Swiss Webster mice that were humanized by oral gavage of human faecal sample obtained from a healthy anonymous donor (American male living in the San Francisco Bay Area, California, age 36, omnivorous diet) as previously described²⁰. Humanized mice closely reconstituted the diversity and phylogenetic make-up of the donor (Pearson's $r=0.96$; Extended Data Fig. 2). The investigators were not blinded to allocation during experiments and outcome assessment. Mice were randomly assigned to two groups and were fed either a high-MAC diet (LabDiet 5010) or a low-MAC diet (Harlan TD.86489). The high-MAC diet is a plant polysaccharide-rich diet in which the MACs come from a diverse source of plants including corn, soybean, wheat, oats, alfalfa and beet. The reported neutral detergent fibre content of the high-MAC diet is 15% by weight. The low-MAC diet is defined as a diet in which carbohydrates are from sucrose (31% by weight), corn starch (31% by weight) and cellulose (5% by weight). The accessibility of cellulose to gut microbiota is known to be extremely low and we have been unable to isolate bacteria from the microbiota that use this substrate²⁰. Mice from the faecal transplant experiment were from the fourth generation of mice from the diet-switching group. Faecal transplants were carried out by gavage using freshly collected faecal samples from fourth-generation control mice consuming a high-MAC diet using a procedure identical to the original humanization²¹. Faecal samples were collected throughout all mouse experiments and stored at -80°C .

16S rRNA amplicon sequencing and analysis. 16S rRNA amplicons were generated for the v4 region from faecal pellets collected and the 16,878,145 Illumina generated sequencing reads were analysed using Qiime 1.8 as described previously¹⁹. Sequencing data underwent quality filtering as described previously and data were rarefied to the lowest number of reads observed in a single sample (28,596 reads)²². OTUs were identified by open-reference picking using the UCLUST algorithm and taxonomy was assigned using the Greengenes 13.8 database. Plots of UniFrac distances are from unweighted analyses and UniFrac distance values are reported as within group versus between groups. Microbiota diversity was measured by Shannon index, which takes into account both overall richness and evenness. High-confidence OTUs were identified using the following criteria: present in at least three mice at the start of the experiment (4 weeks after humanization) and had a

collective abundance of greater than 25 reads. High-confidence sub-OTUs were selected by first filtering out duplicate sequences whose distribution across samples is highly correlated (thresholded at 0.95 dynamical similarity). These duplicates correspond to the same bacterial populations, either as uncommonly frequent errors or as multiple 16S copies within the same bacterium. Second, for each experimental group all sequences whose initial raw abundance was at least 5 reads in at least 3 out of 5 mice in that isolator were selected as high-confidence sub-OTUs. Sub-OTU level analysis was performed as described previously¹⁵.

Glycoside hydrolase profiles. This method was validated using samples for which 16S rRNA profiles and metagenomic data were available (imputed glycoside hydrolase profiles explained 84% of the annotated metagenomic data using a simple linear fit without any model corrections; $P=10^{-28}$)⁵. Glycoside hydrolase imputations were performed by annotating 2,746 reference genomes with glycoside hydrolase families using validated hidden Markov models many of which were derived from conserved domain database models that are capable of identifying several domains within a putative enzyme with increased sensitivity^{23,24}. The glycoside hydrolases from taxonomically assigned communities were then calculated by applying a weighted average of the lowest taxonomic level with representative genomes. To account for the fact that many glycoside hydrolase families have wide ranging functions while each glycoside hydrolase subfamily may possess distinct function, we further clustered the annotated glycoside hydrolases into subfamilies as previously described^{23,25,26}. Fold-changes of glycoside hydrolase family copy numbers were determined by comparing generation-four mice to the mean generation-one glycoside hydrolase profiles with significance testing performed across treatment and control groups.

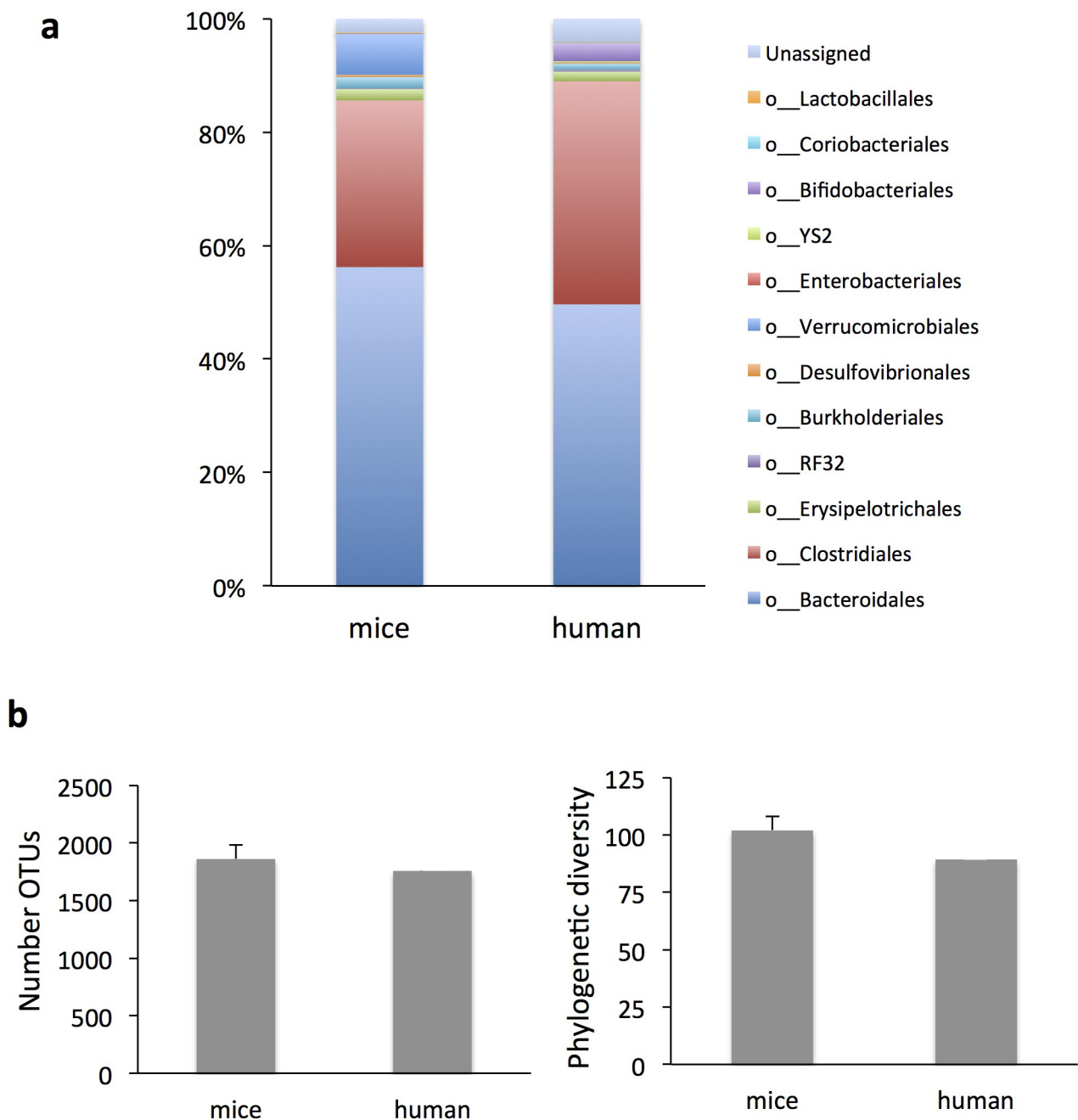
- Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
- Lozupone, C. & Knight, R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228–8235 (2005).
- Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
- Kashyap, P. C. *et al.* Complex interactions among diet, gastrointestinal transit, and gut microbiota in humanized mice. *Gastroenterology* **144**, 967–977 (2013).
- Turnbaugh, P. J. *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* **1**, 6ra14 (2009).
- Bokulich, N. A. *et al.* Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature Methods* **10**, 57–59 (2013).
- Yin, Y. *et al.* dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451 (2012).
- Marchler-Bauer, A. *et al.* CDD: Specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res.* **37**, D205–D210 (2009).
- Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H. & Henrissat, B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol. Biol.* **12**, 186 (2012).
- Stam, M. R., Danchin, E. G. J., Rancurel, C., Coutinho, P. M. & Henrissat, B. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of α -amylase-related proteins. *Protein Eng. Des. Sel.* **19**, 555–562 (2006).



Extended Data Figure 1 | Collating data from studies of the microbiota of hunter-gatherers in Tanzania, agrarians from Malawi and Venezuela, and Westerners from the United States reveals that Western populations have depleted alpha-diversity from birth through childbearing years and are missing bacterial taxa present in the traditional groups.

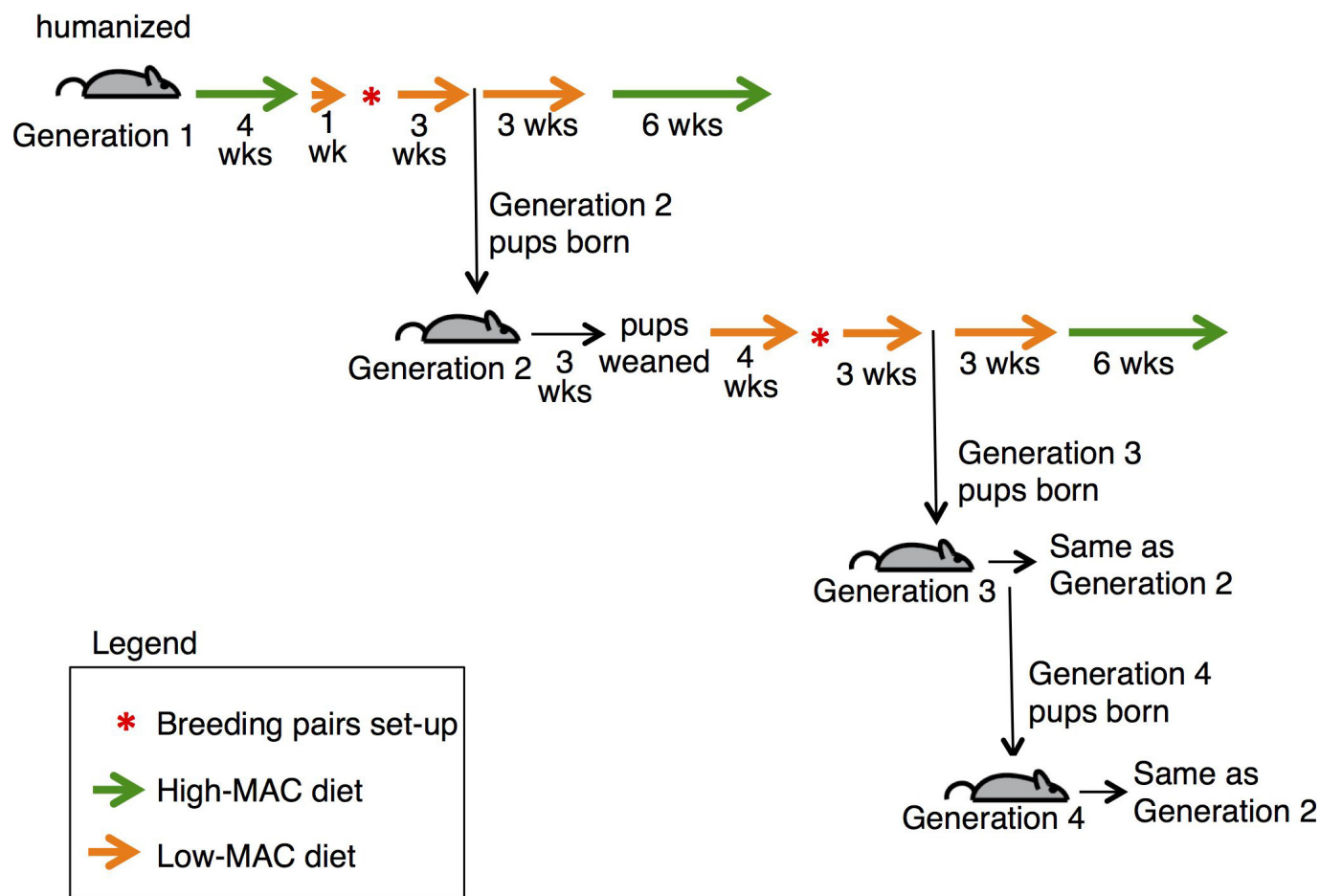
a, Scatterplot of faecal microbiota of individuals plotted by phylogenetic diversity against age of the Hadza hunter-gatherers from Tanzania ($n = 16$, green), agrarians from Malawi ($n = 81$, red) and Venezuela ($n = 78$, purple) and Americans ($n = 213$, blue) **b**, Individuals plotted by unweighted UniFrac PC1 versus phylogenetic diversity. **c**, Individuals plotted by

unweighted UniFrac PC1 versus age. **d**, Line plot of unique OTUs from faecal microbiota across populations (Americans, $n = 315$; Malawi and Venezuela, $n = 213$; Tanzania, $n = 27$). OTUs (x axis; black, present; white, absent) are considered present if represented by $\geq 0.001\%$ of reads within each population. OTUs were sorted along the x axis by their relative abundance in the US and Tanzanian populations and further subdivided by their distributions within a population into tracks (red $> 0.05\%$, yellow $\leq 0.05\%$, and green $\leq 0.01\%$, relative abundance). The opacity of the line is the proportion of that population that meets the criteria for that respective track.



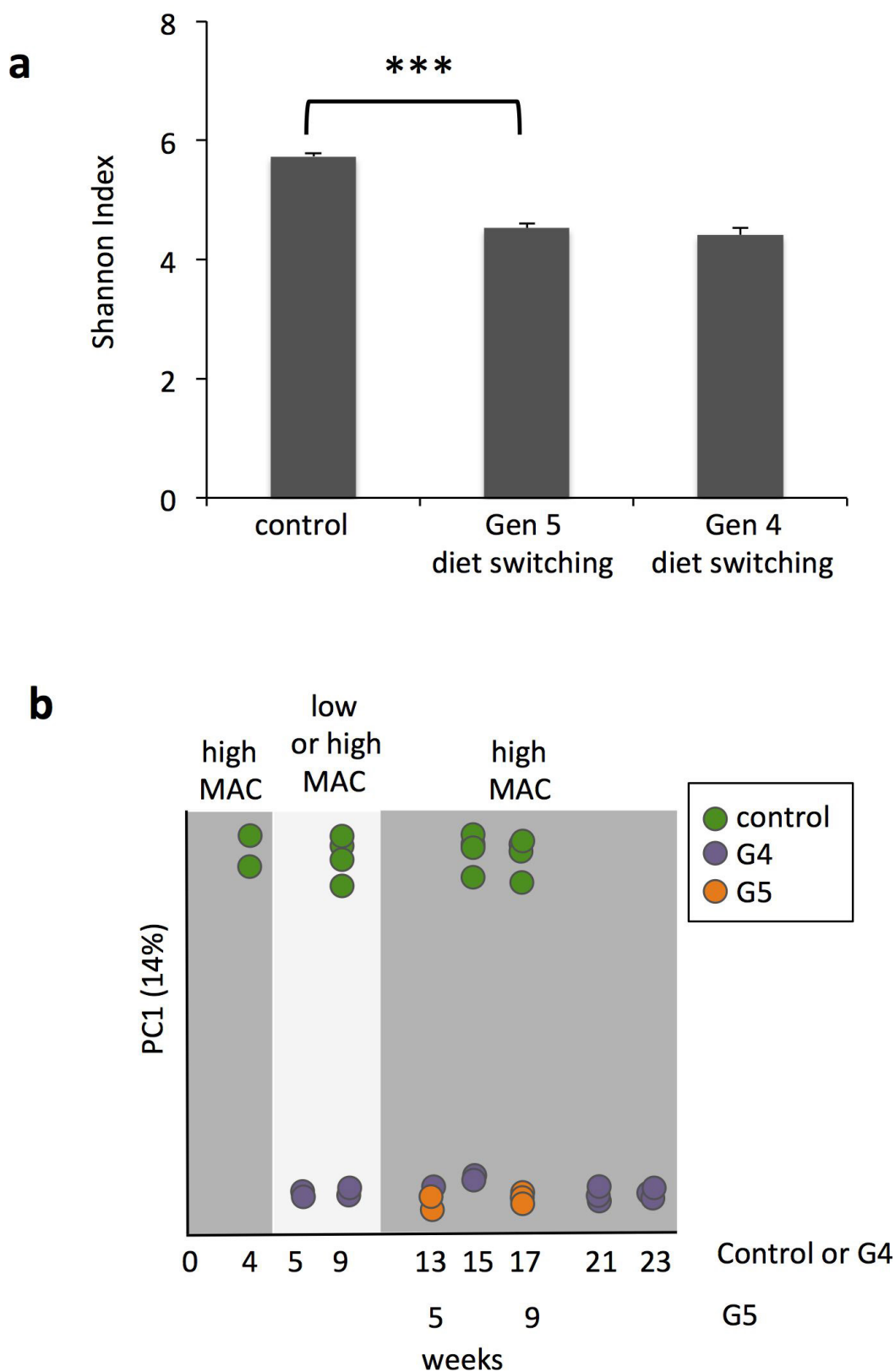
Extended Data Figure 2 | Comparison of human donor and humanized mice. **a**, Taxa summary plot of the relative abundance of taxa from humanized mice faeces (mice) ($n = 10$) and human donor faeces (human)

($n = 1$). **b**, Alpha-diversity of the faecal microbiota from humanized mice (mice) and human donor (human) expressed as number of OTUs (top) and phylogenetic diversity (bottom). Error bars are s.e.m.



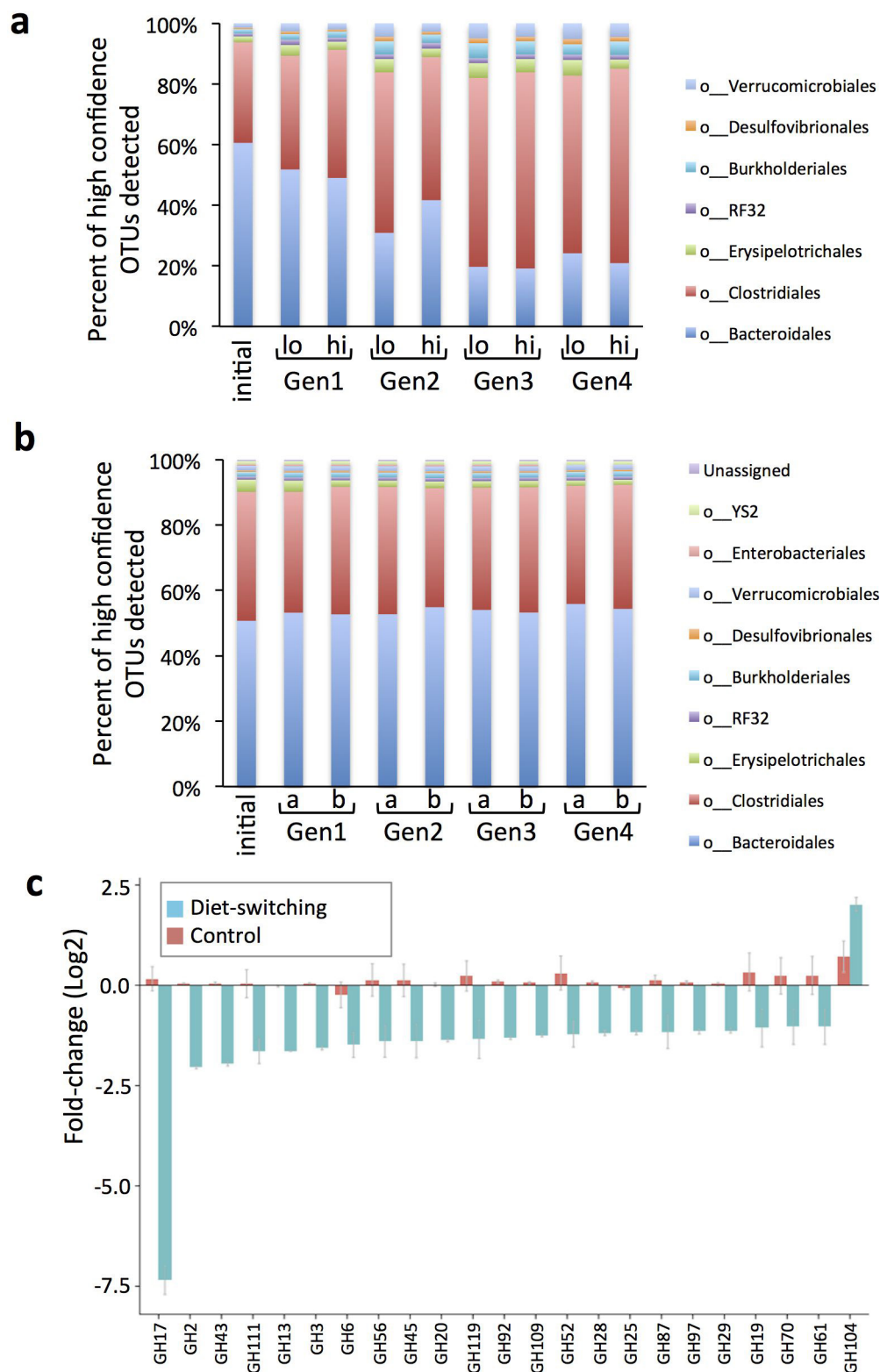
Extended Data Figure 3 | Detailed schematic of multigeneration experiment. Generation one: humanized mice were fed a high-MAC diet for 4 weeks then switched to a low-MAC diet. One week after diet switch, the mice were bred to generate a litter of pups. After three additional weeks on the low-MAC diet, generation-two pups were born and remained in the cage with their mother, suckling for 3 weeks (generation one still consuming the low-MAC diet). After pups were weaned, generation-one mice were returned to the high-MAC diet for 6 weeks. Generation two:

pups were weaned from their mother at 3 weeks old onto a low-MAC diet, which they consumed for 10 weeks. Breeding pairs for generation-two mice were set-up at 7 weeks old. After three additional weeks on the low-MAC diet, generation-three pups were born and remained in the cage with their mother, suckling for 3 weeks (generation-two mice still consuming the low-MAC diet). After pups were weaned, generation-two mice were returned to the high-MAC diet for 6 weeks. Generations three and four followed the same protocol as generation two described above.



Extended Data Figure 4 | Microbiota diversity is not regained after direct weaning the diet-switching group onto the high-MAC diet.
a, Alpha-diversity as measured by Shannon index of faecal microbiota from generation-five mice from the high-MAC-diet control (control) ($n = 6$), generation-five diet-switching group that was weaned directly onto the high-MAC diet (Gen 5 diet switching) ($n = 6$), and generation-four mice from the diet-switching group after weaning and maintenance on the low-MAC diet for 13 weeks and returned to the high-MAC diet for

4 weeks (Gen 4 diet switching) ($n = 5$). Error bars are s.e.m. and P values are from a two-tailed Student's t -test **b**, Principal coordinate analysis of unweighted UniFrac distance for 16S rRNA amplicon profiles from faecal samples collected from first-generation control mice on a high-MAC diet (green), fourth-generation diet-switching mice (purple), and fifth-generation mice from the diet-switching lineage weaned directly onto the high-MAC diet (orange). Control is plotted as weeks post-humanization and generations four and five are plotted as age.

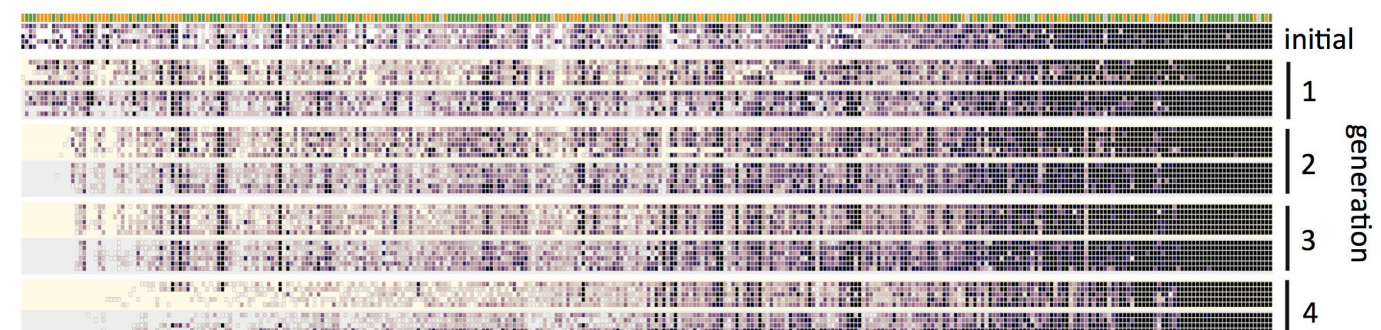
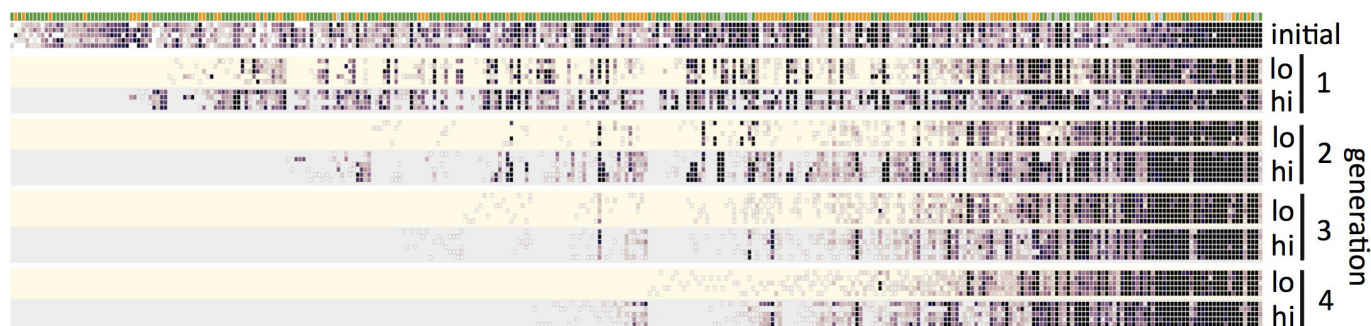


Extended Data Figure 5 | Fraction of high-confidence OTUs from the Clostridiales order increases and from the Bacteroidales order decreases over several generations in the low-MAC-consuming mice.

a, Percentage of high-confidence OTUs, grouped by order, detected in mice faeces over four generations in the diet-switching lineage on the low-MAC diet (lo) and high-MAC diet (hi) ($n = 5$ for Gen 1; $n = 6$ for Gen 2–4). **b**, Percentage of high-confidence OTUs, grouped by order, detected in mice faeces over four generations in the control high-MAC diet lineage

at the equivalent time points to the high-MAC diet (**a**) and low-MAC diet (**b**) of the diet-switching group ($n = 5$ for Gen 1; $n = 6$ for Gen 2–4).

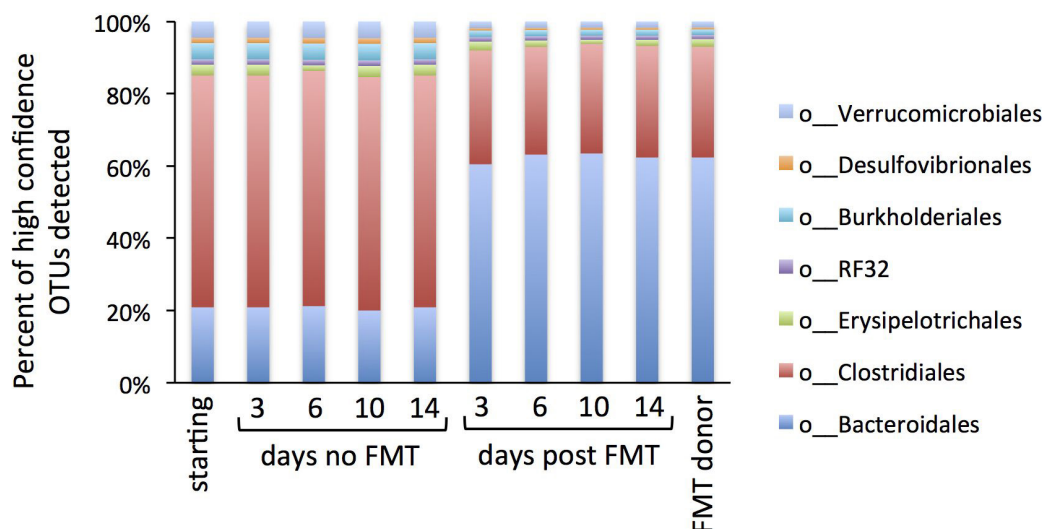
c, Imputed glycoside hydrolase (GH) family members that show significant differences (at least twofold change and $P < 0.05$, Bonferroni-corrected t -test) between generation-four diet-switching mice after 4 weeks on the high-MAC diet (teal) ($n = 5$) and the starting generation-one mice (salmon) ($n = 10$). Error bars depict s.e.m. No glycoside hydrolase families showed significant changes in the control group.



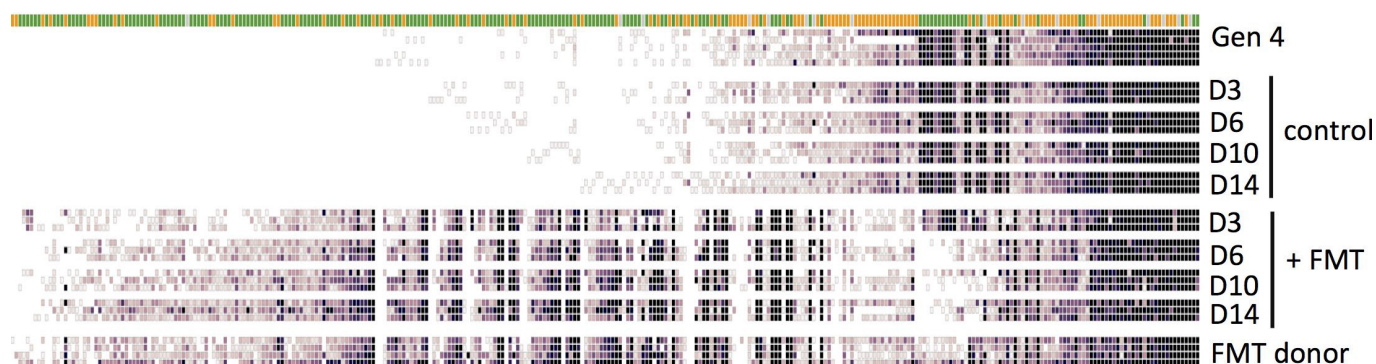
Extended Data Figure 6 | Inefficient inter-generational transfer of taxa driven to low abundance by low dietary MACs. Heat map of abundance of high-confidence sub-OTUs (number of sequencing reads, columns) from faeces of the diet-switching (top) and control (bottom) group. Each row represents an individual mouse faecal microbiota from 4 weeks

post-humanization (initial), while consuming the low-MAC diet (week 9, lo, shaded yellow), and 4 weeks after switching to the high-MAC diet (week 15, hi, shaded grey). Corresponding time points from controls are also shaded. Top row shows the taxonomic assignment for the OTUs plotted: Bacteroidetes are green, Firmicutes are orange, and others are grey.

a



b



Extended Data Figure 7 | Reintroduction of lost taxa and a high-MAC diet restores microbiota diversity and composition with Clostridiales order decreasing and Bacteroidales order increasing in low-MAC-consuming mice that receive a faecal transplant. a, Plot of percentage representation of high-confidence OTUs from generation-four mice faeces in the diet-switching group at day 0 before the FMT (starting) ($n=6$) and then 3–14 days no-FMT control ($n=3$) or post-FMT ($n=3$). FMT donor is plotted on the right. b, Heat map of abundance of high-confidence

sub-OTUs (number of sequencing reads, columns) from the faeces of the diet-switching group at day 0 (Gen 4), days 3–14 that did not receive an FMT (control) ($n=3$ for each day), days 3–14 that received an FMT (+FMT), and the FMT donor. Each row represents an individual mouse faecal microbiota. Top row shows the taxonomic assignment for the OTUs plotted: Bacteroidetes are green, Firmicutes are orange, and others are grey.

Extended Data Table 1 | Nutritional information of mouse diets

Diet Name	Supplier/Product Name	Protein (% by weight)	Carbohydrates (% by weight)	Fat (% by weight)	Fibre (% by weight)	Neutral Detergent Fibre (% by weight)
High-MAC	LabDiet 5010	25	50	5	5	15
Low-MAC	Harlan TD.86489*	18	62	5	5	5**

* Remaining 5% composed of vitamins and minerals

** Exclusively from added cellulose

Extended Data Table 2 | High-confidence OTUs at experiment start

Taxonomy	Control Isolator (# OTUs)	Experimental Isolator (# OTUs)
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__ [Odoribacteraceae]: g__ Butyricimonas: s__	0	1
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Bacteroidaceae: g__Bacteroides: s__	49	48
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Bacteroidaceae: g__Bacteroides: s__fragilis	2	2
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Bacteroidaceae: g__Bacteroides: s__ovatus	2	2
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Bacteroidaceae: g__Bacteroides: s__uniformis	5	6
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Porphyromonadaceae: g__Parabacteroides: s__	3	5
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Porphyromonadaceae: g__Parabacteroides: s__distasonis	2	1
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Porphyromonadaceae: g__Parabacteroides: s__gordonii	1	2
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__Rikenellaceae: g__ : s__	1	1
p__Bacteroidetes: c__Bacteroidia: o__Bacteroidales: f__S24-7: g__ : s__	43	58
p__Cyanobacteria: c__4C0d-2: o__YS2: f__ : g__ : s__	2	0
p__Firmicutes: c__Clostridia: o__Clostridiales: f__ : g__ : s__	3	3
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__ : s__	22	17
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__[Ruminococcus]: s__	3	3
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__[Ruminococcus]: s__gnavus	6	0
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__Blautia: s__	20	17
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__Blautia: s__producta	5	3
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__Coprococcus: s__	6	4
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Lachnospiraceae: g__Dorea: s__	4	3
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Ruminococcaceae: g__ : s__	6	9
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Ruminococcaceae: g__Faecalibacterium: s__prausnitzii	0	4
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Ruminococcaceae: g__Oscillospira: s__	4	3
p__Firmicutes: c__Clostridia: o__Clostridiales: f__Ruminococcaceae: g__Ruminococcus: s__	5	3
p__Firmicutes: c__Erysipelotrichi: o__Erysipelotrichales: f__Erysipelotrichaceae: g__ : s__	4	1
p__Firmicutes: c__Erysipelotrichi: o__Erysipelotrichales: f__Erysipelotrichaceae: g__[Eubacterium]: s__dolichum	2	2
p__Firmicutes: c__Erysipelotrichi: o__Erysipelotrichales: f__Erysipelotrichaceae: g__Coprobacillus: s__	1	1
p__Firmicutes: c__Erysipelotrichi: o__Erysipelotrichales: f__Erysipelotrichaceae: g__Holdemania: s__	1	0
p__Proteobacteria: c__Alphaproteobacteria: o__RF32: f__ : g__ : s__	2	2
p__Proteobacteria: c__Betaproteobacteria: o__Burkholderiales: f__Alcaligenaceae: g__Sutterella: s__	3	3
p__Proteobacteria: c__Deltaproteobacteria: o__Desulfobivibrionales: f__Desulfobivibrionaceae: g__Bilophila: s__	1	1
p__Proteobacteria: c__Gammaproteobacteria: o__Enterobacteriales: f__Enterobacteriaceae: g__ : s__	1	0
p__Verrucomicrobia: c__Verrucomicrobiae: o__Verrucomicrobiales: f__Verrucomicrobiaceae: g__Akkermansia: s__muciniphila	3	3
Unassigned	1	0

Extended Data Table 3 | Shannon index of glycoside hydrolase subfamilies

Comparisons	Gen1 Mouse1	Gen1 Mouse2	Gen1 Mouse3	Gen1 Mouse4	Gen1 Mouse5	Gen4 Mouse1	Gen4 Mouse2	Gen4 Mouse3	Gen4 Mouse4	Gen4 Mouse5	p value	significance sign
Diet-switching	9.39	9.42	9.34	9.35	9.40	9.23	9.21	9.12	9.26	9.17	2.2E-04	**
Control	9.41	9.39	9.45	9.48	9.45	9.42	9.34	9.42	9.41	9.41	0.13	

FOXO1 couples metabolic activity and growth state in the vascular endothelium

Kerstin Wilhelm¹, Katharina Happel¹, Guy Eelen^{2,3}, Sandra Schoors^{2,3}, Mark F. Oellerich¹, Radiance Lim¹, Barbara Zimmermann¹, Irene M. Aspalter⁴, Claudio A. Franco⁵, Thomas Boettger⁶, Thomas Braun⁶, Marcus Fruttiger⁷, Klaus Rajewsky⁸, Charles Keller⁹, Jens C. Brüning¹⁰, Holger Gerhardt^{4,8,11,12,13}, Peter Carmeliet^{2,3} & Michael Potente^{1,14,15}

Endothelial cells (ECs) are plastic cells that can switch between growth states with different bioenergetic and biosynthetic requirements¹. Although quiescent in most healthy tissues, ECs divide and migrate rapidly upon proangiogenic stimulation^{2,3}. Adjusting endothelial metabolism to the growth state is central to normal vessel growth and function^{1,4}, yet it is poorly understood at the molecular level. Here we report that the forkhead box O (FOXO) transcription factor FOXO1 is an essential regulator of vascular growth that couples metabolic and proliferative activities in ECs. Endothelial-restricted deletion of FOXO1 in mice induces a profound increase in EC proliferation that interferes with coordinated sprouting, thereby causing hyperplasia and vessel enlargement. Conversely, forced expression of FOXO1 restricts vascular expansion and leads to vessel thinning and hypobranched. We find that FOXO1 acts as a gatekeeper of endothelial quiescence, which decelerates metabolic activity by reducing glycolysis and mitochondrial respiration. Mechanistically, FOXO1 suppresses signalling by MYC (also known as c-MYC), a powerful driver of anabolic metabolism and growth^{5,6}. MYC ablation impairs glycolysis, mitochondrial function and proliferation of ECs while its EC-specific overexpression fuels these processes. Moreover, restoration of MYC signalling in FOXO1-overexpressing endothelium normalizes metabolic activity and branching behaviour. Our findings identify FOXO1 as a critical rheostat of vascular expansion and define the FOXO1–MYC transcriptional network as a novel metabolic checkpoint during endothelial growth and proliferation.

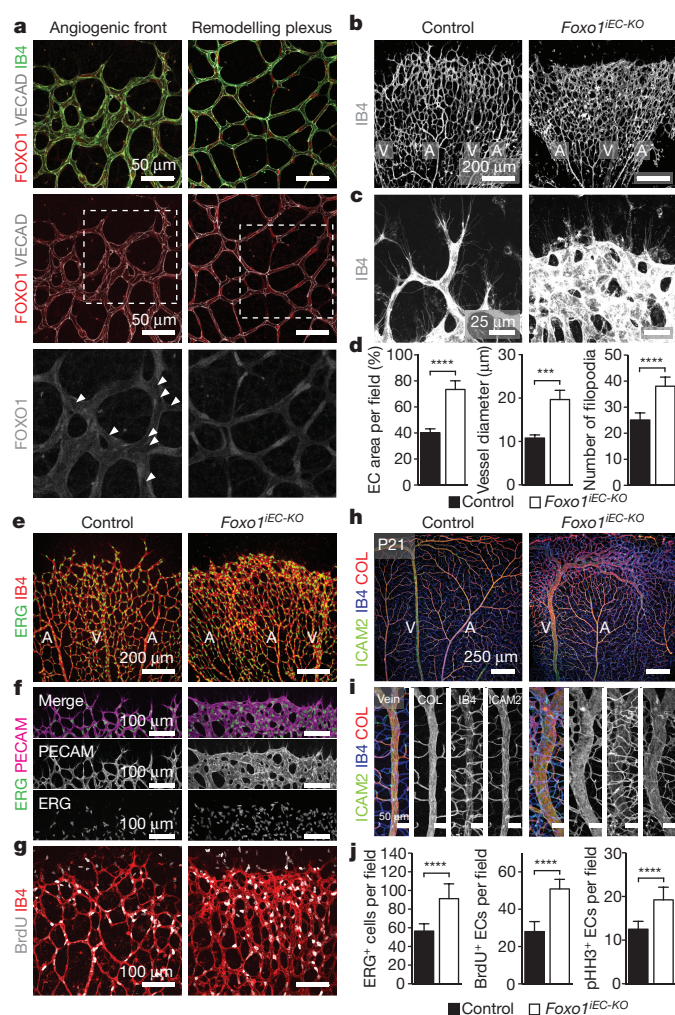
FOXO transcription factors are effectors of the phosphatidylinositol-3-OH kinase (PI(3)K)/AKT pathway that links growth and metabolism^{7,8}. PI(3)K signalling inhibits FOXOs through AKT-mediated phosphorylation leading to their nuclear exclusion^{9,10}. We investigated the role of FOXO1 in ECs, an enriched FOXO family member in the endothelium^{11–15}. To this end, we bred floxed *Foxo1* mice (*Foxo1*^{f/f})¹⁶ with a *Tie2-cre* deleter (Extended Data Fig. 1a), which recombines in endothelial and haematopoietic cells. *Tie2-cre*-mediated deletion of *Foxo1* (*Foxo1*^{EC-KO}) caused defective vascular development and embryonic lethality around embryonic day (E)10.5 (Extended Data Fig. 1b, c)¹⁷, suggesting that endothelial FOXO1 is essential for embryo development.

Immunofluorescence analysis of developing blood vessels in the postnatal retina showed high levels of FOXO1 expression in the endothelium (Fig. 1a). Further examination of the subcellular distribution

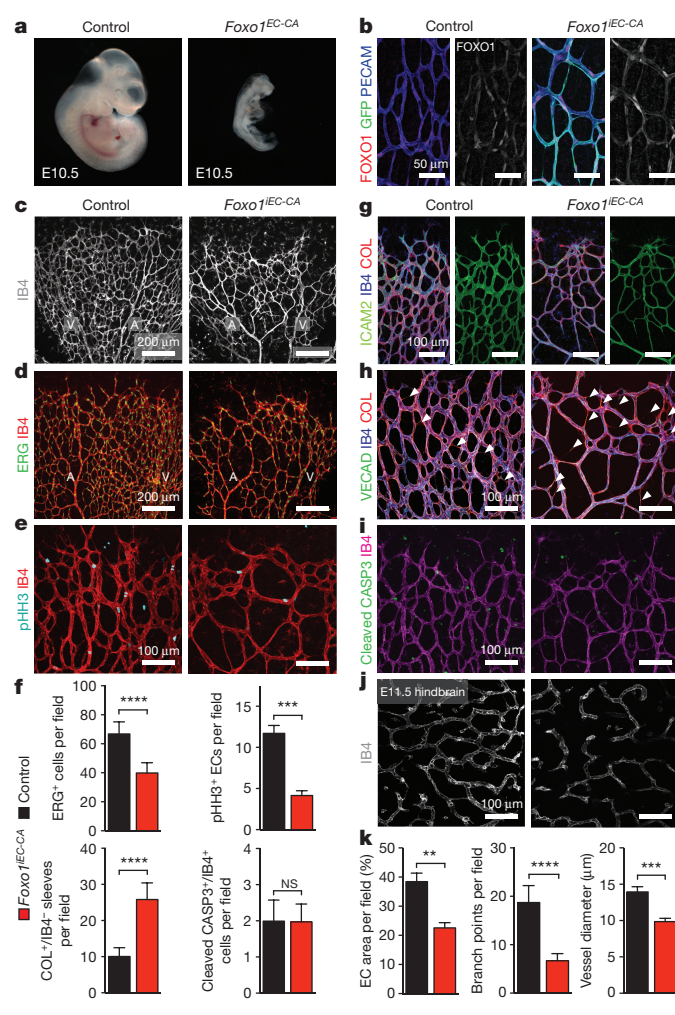
revealed a diffuse nucleocytoplasmic localization of FOXO1 at the angiogenic front, where most of the EC proliferation occurs, but a stronger nuclear pattern in the plexus, where vessels remodel, and endothelial proliferation abates (Fig. 1a). This spatial difference in subcellular localization suggests that FOXO1 is important for governing endothelial growth. To test this, we assessed the impact of *Foxo1* deletion on retinal angiogenesis using the tamoxifen-inducible, endothelial-selective *Pdgfrb-creERT2* line (*Foxo1*^{IEC-KO}). Recombination was monitored with the *Rosa26-mT/mG* (*mTmG*) reporter that expresses green fluorescent protein (GFP) upon Cre-mediated recombination. 4-hydroxy-tamoxifen (4-OHT) treatment resulted in broad GFP expression in ECs as well as extinction of endothelial FOXO1 staining (Extended Data Fig. 1d–f). Endothelial loss of *Foxo1* caused a dense and hyperplastic vasculature and resulted in the inability of ECs to extend proper sprouts (Fig. 1b–f). Instead, ECs grew in clusters leading to vessel enlargement and blunting of the angiogenic front (Fig. 1d, f). Strikingly, numerous filopodial bursts were emanating from the stunted front (Fig. 1c, d), suggesting that FOXO1 deficiency results in uncoordinated vascular growth. Staining *Foxo1*^{IEC-KO} retinas for ERG (marking endothelial nuclei) and VE-cadherin (marking endothelial junctions), revealed an abundance of abnormally aligned ECs, which formed vessels with wide and irregular lumens (Extended Data Fig. 2a–c). Assessment of 5-bromodeoxyuridine (BrdU) incorporation and phospho-histone H3 (pHH3) labelling demonstrated a substantial increase in endothelial proliferation in the *Foxo1*^{IEC-KO} mutants (Fig. 1g, j and Extended Data Fig. 2d), indicating that deregulated proliferation drives this aberrant vessel phenotype. Importantly, the vascular defects of the *Foxo1*^{IEC-KO} mice did not normalize at later stages of development, but showed a persistent increase in endothelial number, density and vessel diameter (Fig. 1h, i and Extended Data Fig. 2e–g). We conclude that FOXO1 is a suppressor of endothelial growth and proliferation, whose inactivation leads to uncontrolled overgrowth.

Next, we determined the consequences of FOXO1 activation in ECs. We used a Cre-inducible gain-of-function allele (*Foxo1*^{CA}) in which the AKT phosphorylation sites are mutated, thus rendering FOXO1 constitutively nuclear (Extended Data Fig. 3a)¹⁸. *Tie2-cre*-mediated expression of this IRES-GFP-coexpressing mutant (*Foxo1*^{EC-CA}) was incompatible with embryo survival beyond E10.5 (Fig. 2a), highlighting the sensitivity of ECs towards changes in FOXO1 status. We then used the *Pdgfrb-creERT2* strain to express *Foxo1*^{CA} in the retinal endothelium (*Foxo1*^{IEC-CA}). Immunofluorescence studies revealed an enriched FOXO1 signal in endothelial nuclei and confirmed the EC-specific

¹Angiogenesis & Metabolism Laboratory, Max Planck Institute for Heart and Lung Research, D-61231 Bad Nauheim, Germany. ²Laboratory of Angiogenesis and Neurovascular Link, Vesalius Research Center, Department of Oncology, University of Leuven, Leuven 3000, Belgium. ³Laboratory of Angiogenesis and Neurovascular Link, Vesalius Research Center, VIB, Leuven 3000, Belgium. ⁴Vascular Biology Laboratory, London Research Institute, Cancer Research UK, London WC2A 3LY, UK. ⁵Vascular Morphogenesis Laboratory, Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisbon 1649-028, Portugal. ⁶Department of Cardiac Development and Remodeling, Max Planck Institute for Heart and Lung Research, D-61231 Bad Nauheim, Germany. ⁷UCL Institute of Ophthalmology, University College London, London EC1V 9EL, UK. ⁸Max Delbrück Center for Molecular Medicine (MDC), D-13125 Berlin, Germany. ⁹Children's Cancer Therapy Development Institute, Beaverton, Oregon 97005, USA. ¹⁰Max Planck Institute for Metabolism Research, Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD) and Center of Molecular Medicine Cologne (CMCC), Center for Endocrinology, Diabetes and Preventive Medicine (CEDP), University of Cologne, D-50931 Cologne, Germany. ¹¹Vascular Patterning Laboratory, Vesalius Research Center, VIB and University of Leuven, Leuven 3000, Belgium. ¹²DZHK (German Center for Cardiovascular Research), partner site Berlin, D-13347 Berlin, Germany. ¹³Berlin Institute of Health (BIH), D-10117 Berlin, Germany. ¹⁴International Institute of Molecular and Cell Biology, 02-109 Warsaw, Poland. ¹⁵DZHK (German Center for Cardiovascular Research), partner site Frankfurt Rhine-Main, D-13347 Berlin, Germany.



expression of GFP (Fig. 2b and Extended Data Fig. 3b, c, g). Forced activation of FOXO1 led to a sparse and hyperpruned vascular network that contained fewer ECs (Fig. 2c, d, f–h and Extended Data Fig. 3g). These retinal vessels established a lumen but were thinner and extended fewer filopodia (Fig. 2g, h and Extended Data Fig. 3d). Staining for BrdU incorporation and pHH3 revealed a reduction in EC proliferation in *Foxo1*^{IEC-CA} mice while endothelial apoptosis was not altered (Fig. 2e, f, i and Extended Data Fig. 3e, g). Similar phenotypes were observed in the hindbrain vasculature (Fig. 2j, k and Extended Data Fig. 4a–c), indicating that FOXO1 is a critical driver of endothelial quiescence. To examine this further, we analysed mosaic retinas of *mTmG*-coexpressing control and *Foxo1*^{IEC-CA} mice, in which the majority of ECs are unrecombined. Compared to controls, *Foxo1*^{IEC-CA} mice showed an impaired propagation of GFP-positive ECs in the retinal



plexus (Extended Data Fig. 4d–f), arguing that the proquiescent activity of FOXO1 is cell autonomous.

We next assessed whether FOXO1 regulates endothelial metabolism. Since ECs rely on glycolysis for vessel branching⁴, we first studied the effects of FOXO1 on this metabolic pathway. Transduction of human umbilical vein endothelial cells (HUVECs) with a FOXO1^{CA}-encoding adenovirus (AdFOXO1^{CA}) led to a robust reduction in glycolysis as evidenced by a reduction in extracellular acidification rate (ECAR), glucose uptake, glycolytic flux and lactate production (Fig. 3a–d and Extended Data Fig. 5a). This metabolic phenotype correlates with the reduced proliferation in FOXO1^{CA}-expressing ECs and raises a question as to whether FOXO1 promotes mitochondrial oxidative phosphorylation. Surprisingly, FOXO1 did not stimulate but instead diminished oxidative metabolism as indicated by a decline in oxygen consumption in AdFOXO1^{CA}-expressing HUVECs (Fig. 3e). Moreover, reactive oxygen species (ROS) formation and ATP levels were decreased (Fig. 3f and

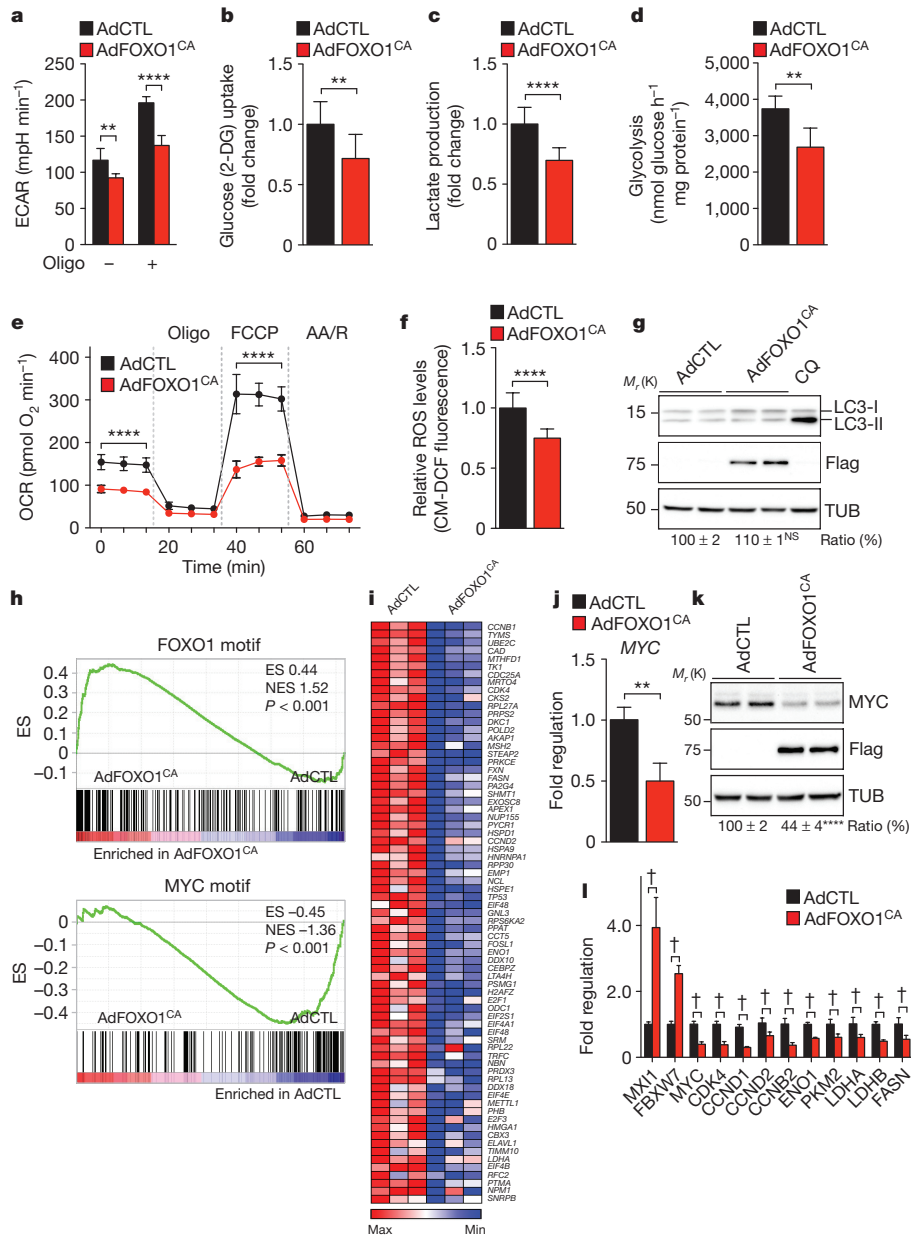


Figure 3 | FOXO1 slows endothelial metabolic activity and suppresses MYC signalling. **a**, ECAR in ECs treated with or without oligomycin (Oligo) showing reduced basal and maximal glycolytic activity in AdFOXO1^{CA}- compared to control adenovirus (AdCTL)-transduced HUVECs (*n* = 6). **b–d**, Reduced 2-deoxy-D-glucose (2-DG) uptake (**b**; *n* = 13), relative lactate production (**c**; *n* = 10) and glycolytic flux (**d**; *n* = 4) in FOXO1^{CA}-expressing ECs. **e**, Oxygen consumption rates (OCR) in control and FOXO1^{CA}-overexpressing ECs (*n* = 5) under basal conditions and in response to oligomycin, fluoro-carbonyl cyanide phenylhydrazone (FCCP) and antimycin A (AA)/rotenone (R). Data represent mean ± s.d. Two-way analysis of variance (ANOVA) with Bonferroni's multiple comparison test. **f**, Relative ROS levels in AdCTL- or AdFOXO1^{CA}-transduced ECs (*n* = 7). CM-DCF, 5-(and-6)-chloromethyl-2',7'-dichlorodihydrofluorescein diacetate. **g**, LC3 western blot analysis

Extended Data Fig. 5b). Importantly, FOXO1 did not induce endothelial apoptosis, senescence, autophagy or energy distress under the same experimental conditions (Fig. 3g and Extended Data Fig. 5c–h). Together, our data indicate that FOXO1 adapts metabolic activity to the lower requirements of the quiescent endothelium.

To gain insight into the underlying mechanisms for this adaptability, we performed transcriptome analysis of FOXO1^{CA}- and

showing that overexpression of the Flag-tagged FOXO1^{CA} does not induce autophagy in ECs. CQ, chloroquine; TUB, tubulin. Densitometric quantifications are shown below the lanes ($n = 10$). **h**, GSEA of the FOXO1 (AAACAA) or MYC (CACGTG) DNA-binding element gene sets in AdFOXO1^{CA}- or AdCTL-transduced ECs. ES, enrichment score; NES, normalized enrichment score. **i**, Heat map of downregulated MYC signature genes in FOXO1^{CA}-overexpressing ECs ($n = 3$). **j, k**, Analysis of MYC expression by microarray (**j**) and immunoblot (**k**) in FOXO1^{CA}-Flag-overexpressing endothelium. **j**, $n = 6$; **k**, $n = 10$. **l**, Quantitative polymerase chain reaction (qPCR) expression analysis of FOXO1^{CA}-regulated genes involved in MYC signalling ($n \geq 3$). **a–d, f, g, j–l**, Data represent mean \pm s.d., two-tailed unpaired *t*-test. ** $P < 0.01$; **** $P < 0.0001$; † $P < 0.01$; NS, not significant.

GFP-transduced HUVECs. Gene set enrichment analysis (GSEA) revealed an enrichment of the FOXO1 DNA-binding elements in genes induced by FOXO1, while the MYC DNA-binding motif was highly enriched in the repressed genes (Fig. 3h and Extended Data Fig. 6a, b). Moreover, MYC target gene signatures were downregulated in the FOXO1 transcriptome (Fig. 3i and Extended Data Fig. 6c). Since MYC is a powerful driver of glycolysis, mitochondrial metabolism and

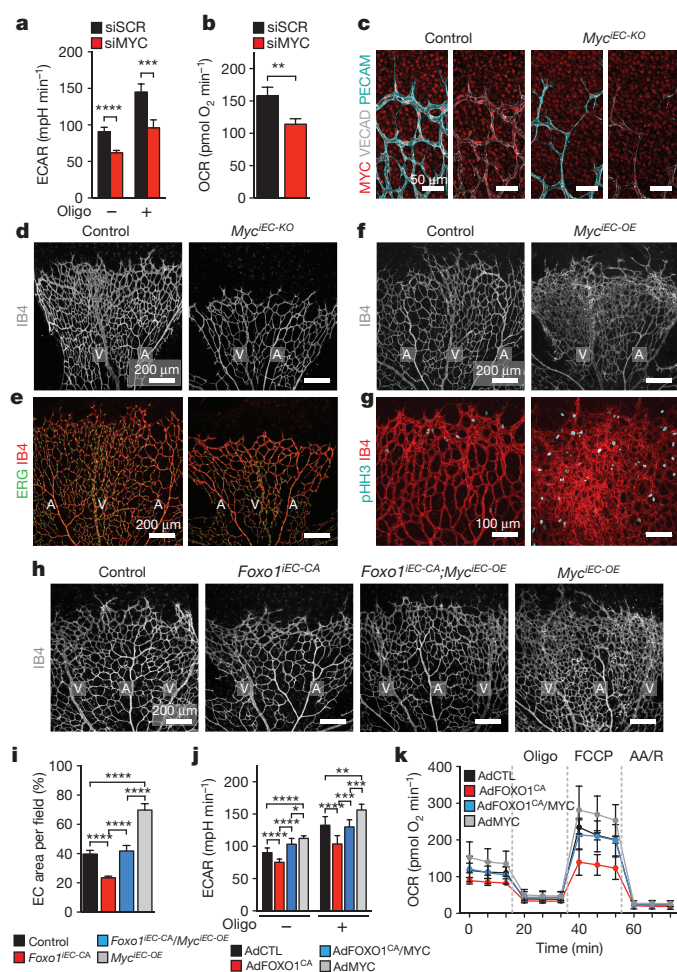


Figure 4 | MYC is a critical component of FOXO1 signalling in ECs. **a, b**, ECAR (**a**) and OCR (**b**) in MYC siRNA- (siMYC) or scrambled siRNA- (siSCR) transfected ECs (ECAR: $n = 5$; OCR: $n = 5$). Data represent mean \pm s.d., two-tailed unpaired t -test. **c**, Staining for MYC, VECAD and PECAM in retinas of *Myc*^{IEC-KO} and control mice. **d, e**, Images of IB4- (**d**) and IB4- and ERG- (**e**) stained P5 retinas of control and *Myc*^{IEC-KO} mice. **f, g**, Images of IB4- (**f**) and pHH3- and IB4- (**g**) stained P5 retinal vessels in *Myc*^{IEC-OE} and control mice. **h, i**, Representative images (**h**) and quantification (**i**) of IB4-stained P5 retinas in control, *Foxo1*^{IEC-CA}, *Myc*^{IEC-OE} and *Foxo1*^{IEC-CA}/*Myc*^{IEC-OE} double mutants ($n \geq 6$). **j, k**, ECAR (**j**) and OCR (**k**) in AdCTL, AdFOXO1^{CA}, AdFOXO1^{CA}/AdMYC and AdMYC-transduced HUVECs showing restoration of metabolic activity in FOXO1^{CA}/MYC co-expressing ECs (ECAR: $n = 8$; OCR: $n \geq 3$). **i–k**, Data represent mean \pm s.d., one-way ANOVA with Bonferroni's multiple comparison post-hoc test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

growth⁵, FOXO1 might antagonize endothelial MYC signalling. In line with this, overexpression of FOXO1^{CA} suppressed MYC expression, whereas FOXO1 depletion enhanced MYC levels, both in HUVECs and in ECs derived from the mutant mice (Fig. 3j–l and Extended Data Fig. 6d–h). Immunofluorescence studies in *Foxo1*^{IEC-CA} retinas confirmed these findings and showed a suppression of endothelial MYC expression (Extended Data Fig. 3f). Accordingly, numerous genes that are induced by MYC were downregulated in FOXO1^{CA}-overexpressing HUVECs, including genes involved in cell metabolism and cell cycle progression (Fig. 3i, l). This regulation is in line with the repression of MYC by FOXOs in cancer cells^{19–23} and points to MYC as a crucial effector of FOXO1 in the coordination of endothelial metabolism and growth.

Remarkably, FOXO1 also induced the expression of negative regulators of MYC signalling including *MXI1*, an antagonist of MYC

transcriptional activity⁶, and *FBXW7*, an E3 ubiquitin ligase that targets MYC for proteasomal degradation⁶ (Fig. 3l and Extended Data Fig. 7a). Consistent with these findings, MYC protein stability was decreased in FOXO1^{CA}-expressing ECs and co-treatment with the proteasomal inhibitor MG132 partially restored MYC protein levels (Extended Data Fig. 7b, c). Knockdown of *MXI1*, on the other hand, did not affect the FOXO1-induced repression of MYC but attenuated the ability of FOXO1 to downregulate MYC target genes (Extended Data Fig. 7d–f). These data are in accordance with the function of *MXI1* as a negative regulator of MYC activity^{6,19,20}, and suggest that FOXO1 intersects with MYC signalling at different levels.

To explore further the role of MYC in ECs, we profiled the transcriptome of MYC short interfering RNA (siRNA)-transfected (siMYC) HUVECs. This analysis showed a suppression of MYC signature genes involved in metabolic processes and cell cycle progression and validated the regulation of predicted MYC targets that were repressed by FOXO1 (Extended Data Fig. 8a–g). Bioenergetic analysis revealed that MYC deficiency attenuated glycolysis and mitochondrial respiration, whereas adenoviral overexpression of MYC (AdMYC) stimulated these metabolic activities (Fig. 4a, b and Extended Data Fig. 9e). Conditional deletion of *Myc* (*Myc*^{fl/fl})²⁴ in mice using the *Pdgfb-creERT2* deleter impaired vascular expansion and led to a thinned and poorly branched vasculature with reduced EC proliferation (Fig. 4c–e and Extended Data Fig. 9a–d). These phenotypes resemble the vascular defects in *Foxo1*^{IEC-CA} mutant mice and imply that MYC is a central component of endothelial FOXO1 signalling. To test this directly, we attempted to rescue the endothelial phenotypes imposed by FOXO1 activation by restoring MYC signalling with a Cre-inducible *Myc* overexpressor allele (*Myc*^{OE})²⁵. *Pdgfb-creERT2*-induced overexpression of MYC caused sustained vascular overgrowth and led to a profound increase in EC number, proliferation and vessel density (Fig. 4f, g and Extended Data Fig. 9f–k). We then combined the *Myc*^{OE}, *Foxo1*^{CA} and *Pdgfb-creERT2* alleles to generate endothelial-specific double mutants. Remarkably, re-expression of MYC in ECs of *Foxo1*^{IEC-CA} mice normalized the hypobranching and hypocellular vascular phenotype caused by FOXO1 activation (Fig. 4h, i and Extended Data Fig. 10a, b). Moreover, coexpression of MYC and FOXO1^{CA} in HUVECs restored glycolysis, mitochondrial respiration and ROS formation (Fig. 4j, k and Extended Data Fig. 10c), indicating that regulation of MYC signalling by FOXO1 is critical for the coordination of endothelial metabolism and growth.

This study identifies FOXO1 as a critical checkpoint of endothelial growth that restricts vascular expansion. Our data suggest that FOXO1 promotes endothelial quiescence by antagonizing MYC, which leads to a coordinated reduction in the proliferative and metabolic activity of ECs. The FOXO1-induced deceleration of metabolic activity might not only enforce quiescence but also support endothelial function. For instance, by lowering metabolism, ECs will consume less energetic fuel for their homeostatic needs, thereby ensuring efficient nutrient and oxygen delivery. Reducing metabolic activity might also contribute to endothelial redox balance. ECs are long-lived cells that need to protect themselves against oxidative damage exerted by high oxygen levels in the bloodstream. The FOXO1-induced reduction in oxidative metabolism might thus be a mechanism to minimize the production of mitochondria-derived ROS, thereby conferring protection against the high-oxygen environment. Such a role of FOXO1 in endothelial metabolism aligns with the broader function of FOXOs in mediating oxidative stress resistance^{26–30}, and might also explain why ECs are exquisitely sensitive to a change in FOXO1 status. It will be interesting to determine how endothelial FOXO1 is regulated *in vivo* and how deregulation contributes to disease.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 August 2014; accepted 27 November 2015.

Published online 6 January 2016.

1. Ghesquière, B., Wong, B. W., Kuchnio, A. & Carmeliet, P. Metabolism of stromal and immune cells in health and disease. *Nature* **511**, 167–176 (2014).
2. Adams, R. H. & Alitalo, K. Molecular regulation of angiogenesis and lymphangiogenesis. *Nature Rev. Mol. Cell Biol.* **8**, 464–478 (2007).
3. Potente, M., Gerhardt, H. & Carmeliet, P. Basic and therapeutic aspects of angiogenesis. *Cell* **146**, 873–887 (2011).
4. De Bock, K., Georgiadou, M. & Carmeliet, P. Role of endothelial cell metabolism in vessel sprouting. *Cell Metab.* **18**, 634–647 (2013).
5. Dang, C. V. MYC, metabolism, cell growth, and tumorigenesis. *Cold Spring Harb. Perspect. Med.* **3**, a014217 (2013).
6. Adhikary, S. & Eilers, M. Transcriptional regulation and transformation by Myc proteins. *Nature Rev. Mol. Cell Biol.* **6**, 635–645 (2005).
7. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).
8. Ward, P. S. & Thompson, C. B. Metabolic reprogramming: a cancer hallmark even Warburg did not anticipate. *Cancer Cell* **21**, 297–308 (2012).
9. Salihi, D. A. & Brunet, A. FoxO transcription factors in the maintenance of cellular homeostasis during aging. *Curr. Opin. Cell Biol.* **20**, 126–136 (2008).
10. Eijkelenboom, A. & Burgering, B. M. FOXOs: signalling integrators for homeostasis maintenance. *Nature Rev. Mol. Cell Biol.* **14**, 83–97 (2013).
11. Daly, C. *et al.* Angiopoietin-1 modulates endothelial cell function and gene expression via the transcription factor FKHR (FOXO1). *Genes Dev.* **18**, 1060–1071 (2004).
12. Potente, M. *et al.* Involvement of Foxo transcription factors in angiogenesis and postnatal neovascularization. *J. Clin. Invest.* **115**, 2382–2392 (2005).
13. Paik, J. H. *et al.* FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell* **128**, 309–323 (2007).
14. Taddei, A. *et al.* Endothelial adherens junctions control tight junctions by VE-cadherin-mediated upregulation of claudin-5. *Nature Cell Biol.* **10**, 923–934 (2008).
15. Tsuchiya, K. *et al.* FoxOs integrate pleiotropic actions of insulin in vascular endothelium to protect mice from atherosclerosis. *Cell Metab.* **15**, 372–381 (2012).
16. Keller, C. *et al.* Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev.* **18**, 2614–2626 (2004).
17. Sengupta, A., Chakraborty, S., Paik, J., Yutzy, K. E. & Evans-Anderson, H. J. FoxO1 is required in endothelial but not myocardial cell lineages during cardiovascular development. *Dev. Dyn.* **241**, 803–813 (2012).
18. Stöhr, O. *et al.* Insulin receptor signaling mediates APP processing and β -amyloid accumulation without altering survival in a transgenic mouse model of Alzheimer's disease. *Age (Dordr)* **35**, 83–101 (2013).
19. Delpuech, O. *et al.* Induction of Mxi1-SR α by FOXO3a contributes to repression of Myc-dependent gene expression. *Mol. Cell Biol.* **27**, 4917–4930 (2007).
20. Gan, B. *et al.* FoxOs enforce a progression checkpoint to constrain mTORC1-activated renal tumorigenesis. *Cancer Cell* **18**, 472–484 (2010).
21. Jensen, K. S. *et al.* FoxO3a promotes metabolic adaptation to hypoxia by antagonizing Myc function. *EMBO J.* **30**, 4554–4570 (2011).
22. Kress, T. R. *et al.* The MK5/PRAK kinase and Myc form a negative feedback loop that is disrupted during colorectal tumorigenesis. *Mol. Cell* **41**, 445–457 (2011).
23. Ferber, E. C. *et al.* FOXO3a regulates reactive oxygen metabolism by inhibiting mitochondrial gene expression. *Cell Death Differ.* **19**, 968–979 (2012).
24. de Alboran, I. M. *et al.* Analysis of C-MYC function in normal cells via conditional gene-targeted mutation. *Immunity* **14**, 45–55 (2001).
25. Sander, S. *et al.* Synergy between PI3K signaling and MYC in Burkitt lymphomagenesis. *Cancer Cell* **22**, 167–179 (2012).
26. Kops, G. J. *et al.* Forkhead transcription factor FOXO3a protects quiescent cells from oxidative stress. *Nature* **419**, 316–321 (2002).
27. Nemoto, S. & Finkel, T. Redox regulation of forkhead proteins through a p66shc-dependent signaling pathway. *Science* **295**, 2450–2452 (2002).
28. Tothova, Z. *et al.* FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* **128**, 325–339 (2007).
29. Miyamoto, K. *et al.* Foxo3a is essential for maintenance of the hematopoietic stem cell pool. *Cell Stem Cell* **1**, 101–112 (2007).
30. Yeo, H. *et al.* FoxO3 coordinates metabolic pathways to maintain redox balance in neural stem cells. *EMBO J.* **32**, 2589–2602 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank F. W. Alt and I. Aifantis for *Myc^{fl/fl}* mice, and T. Enders, J. Sperling and K. Wilson for assistance with the mouse colony. The research of M.P. is supported by the Max Planck Society, the European Research Council (ERC) Starting Grant ANGIOMET (311546), the Deutsche Forschungsgemeinschaft (SFB 834), the Excellence Cluster Cardiopulmonary System (EXC 147/1), the Cluster of Excellence Macromolecular Complexes (EXC 115), the Foundation Leducq Transatlantic Network (ARTEMIS), the LOEWE grant Ub-Net, and the European Molecular Biology Organization Young Investigator Programme. S.S. was supported by a PhD fellowship of the Belgian Science Policy (IWT) and I.M.A. is a recipient of a DOC-fORTE fellowship of the Austrian Academy of Sciences. C.A.F. was supported by a European Union FP7 Marie Curie Post-doctoral Fellowship and a FCT grant (IF/00412/2012). The research of K.R. is supported by an ERC Advanced Grant (268921) and C.K. received support from the National Institutes of Health K08CA090438. The research of H.G. is supported by Cancer Research UK, an ERC consolidator grant (REshape), the Foundation Leducq Transatlantic Network (ARTEMIS), the Lister Institute of Preventive Medicine and the British Council under the BIRAX Initiative. The work of P.C. is funded by long-term structural funding: Methusalem Funding by the Flemish Government, ERC Advanced Research Grant (269073), and FWO G.0834.13N from the Flanders Science Fund.

Author Contributions K.W., K.H., G.E., S.S., M.F.O., R.L., B.Z., I.M.A., C.A.F., T.Bo. and M.P. performed experiments. K.W., K.H., G.E., S.S., M.F.O., R.L., B.Z., I.M.A., C.A.F., T.Bo., H.G., P.C. and M.P. analysed data. H.G., P.C. and M.P. guided research. T.Br., M.F., K.R., C.K., J.C.B. and H.G. provided essential reagents and protocols. K.H. contributed to manuscript writing and preparation. K.W., H.G., P.C. and M.P. wrote the paper. M.P. conceived and directed the study. All authors discussed the results and commented on the manuscript.

Author Information Data from microarray analysis have been deposited in the ArrayExpress repository under the accessions E-MTAB-4023 and E-MTAB-4025. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.P. (michael.potente@mpi-bn.mpg.de).

METHODS

Animals and genetic experiments. All conditional *Foxo1*- and *Myc*-mutant mice were on a C57BL/6 genetic background and generated as described^{16,18,24,25}. For constitutive Cre-mediated recombination in ECs, *Foxo1^{fl/fl}* or *Rosa26-Foxo1^{CA}* mice were bred with *Tie2-cre* transgenic mice³¹. To avoid recombination in the female germline, only *Tie2-cre*-positive male mice were used for intercrossing. Embryos were collected from *cre*-negative females at the indicated time points and genotyping was performed from isolated yolk sacs. For inducible Cre-mediated recombination in ECs, floxed mice were bred with transgenic mice expressing the tamoxifen-inducible, *Pdgfrb* promoter-driven *creERT2* recombinase³². The degree of Cre-mediated recombination was assessed with the double-fluorescent Cre-reporter *Rosa26-mT/mG³³* allele, which was crossed into the respective mutant mice. For the analysis of angiogenesis in the postnatal mouse retina, Cre-mediated recombination was induced in newborn mice by intraperitoneal (i.p.) injections of 25 μ l 4-hydroxy-tamoxifen (4-OHT; 2 mg ml⁻¹; Sigma-Aldrich) from postnatal day (P)1 to P4. Eyes were harvested at P5 or P21 for further analysis. In mosaic recombination experiments, 4-OHT (20 μ g g⁻¹ body weight of 0.02 mg ml⁻¹) was injected i.p. at P3 and eyes were collected at P5. To induce Cre-mediated recombination in mouse embryos, 100 μ l of 4-OHT (10 mg ml⁻¹) was injected i.p. into pregnant females from embryonic day (E)8.5 to E10.5. Embryos were harvested at E11.5 for the analysis of angiogenesis in the embryonic hindbrain. The *Rosa26-Foxo1^{CA}*, *Rosa26-Myc* and *Rosa26-mT/mG* alleles were kept heterozygous for the respective transgene in all experimental studies. Apart from the mosaic studies, control animals were littermate animals without *cre* expression. Male and female mice were used for the analysis, which were maintained under specific pathogen-free conditions. Experiments involving animals were conducted in accordance with institutional guidelines and laws, following protocols approved by local animal ethics committees and authorities (Regierungspresidium Darmstadt).

Immunohistochemistry of mouse tissues. To analyse blood vessel growth in the postnatal retina, whole mouse eyes were fixed in 4% paraformaldehyde (PFA) on ice for 1 h. Eyes were washed in PBS before the retinas were dissected and partially cut into four leaflets. After blocking/permeabilization in 2% goat serum (Vector Laboratories), 1% BSA and 0.5% Triton X-100 (in PBS) for 1 h at room temperature, the retinas were incubated at 4°C overnight in incubation buffer containing 1% goat serum, 0.5% BSA and 0.25% Triton X-100 (in PBS) and the primary antibody. Primary antibodies against the following proteins were used: cleaved caspase 3 (Cell Signaling Technology, #9664, 1:100), collagen IV (AbD Serotec, #2150-1470, 1:400), ERG 1/2/3 (Abcam, #ab92513, 1:200), FOXO1 (Cell Signaling Technology, #2880, 1:100), GFP (Invitrogen, #A21311, 1:100), ICAM2 (BD Biosciences, #553326, 1:200), MYC (Millipore, 06-340, 1:100), PECAM-1 (R&D Systems, AF3628, 1:400), phospho-histone H3 (Chemicon, #06-570, 1:100), TER119 (BD Biosciences, #553670, 1:100), and VE-cadherin (BD Biosciences, #555289, 1:25). After four washes with 0.1% Triton X-100 in PBS (PBST), retinas were incubated with Alexa-Fluor 488-, Alexa-Fluor 555- or Alexa-Fluor 647-conjugated secondary antibodies (Invitrogen, 1:400) for 2 h at room temperature. For staining ECs with isolectin B4 (IB4), retinas were washed with PBLEC buffer (1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂ and 1% Triton X-100 in PBS) and incubated with biotinylated IB4 (*Griffonia simplicifolia*, #B1205, Vector Laboratories, 1:100) diluted in PBLEC buffer. After washing, retinas were incubated in Alexa-Fluor-coupled streptavidin (Invitrogen, #S21374, 1:200) for 2 h at room temperature. For nuclear counterstain, retinas were incubated with 4',6-diamidino-2-phenylindole (DAPI; Sigma Aldrich, #D9542, 1:1,000) for 15 min following washes with PBST and PBS. The labelling of proliferating cells with BrdU was performed in P5 pups. In brief, 50 mg kg⁻¹ of BrdU (Invitrogen, #B23151) per pup was injected i.p. 3 h before they were killed. Retinas were fixed for 2 h in 4% PFA and then incubated for 1 h in 65°C warm formaldehyde, followed by an incubation of 30 min in 2 N HCl. Afterwards retinas were washed twice with 0.1 M Tris-HCl (pH 8) and then blocked in 1% BSA, 0.5% Tween 20 in PBS and incubated overnight at 4°C with a mouse anti-BrdU antibody (BD Biosciences, #347580, 1:50). The detection was performed with Alexa-Fluor-488 anti-mouse secondary antibody (Invitrogen, A21202, 1:400). After the BrdU staining, retinas were processed for the IB4 staining as described earlier. The dissection of the embryonic hindbrain was performed as described³⁴. After overnight fixation in 4% PFA, dissected hindbrains were incubated in a blocking solution containing 10% serum, 1% BSA and 0.5% Triton X-100 in PBS at 4°C. After washes with PBS, hindbrains were incubated for 1 h in PBLEC buffer before the overnight incubation with Alexa-Fluor-conjugated IB4 (Invitrogen, #I21411, 1:100 in PBLEC) at 4°C. Hindbrains were washed with PBS and stained with DAPI. Retinas and embryonic hindbrains were flat-mounted with Vectashield (Vector Laboratories) and examined by confocal laser microscopy (Leica TCS SP5 or SP8). Immunostainings were carried out in tissues from littermates and processed under the same conditions.

Immunohistochemistry of cell cultures. HUVECs were seeded on glass-bottom culture dishes (Mattek) and cultured at 37°C and 5% CO₂. To detect autophagy, cells were washed and fixed with 4% PFA for 20 min at room temperature.

Permeabilization was performed in 1% BSA, 10% donkey serum and 0.5% Tween-20 in PBS. Cells were stained for anti-LC3A/B (Cell Signaling Technology, #12741, 1:400), Phalloidin-TRITC (Sigma Aldrich, #P1951, 1:500) and DAPI in incubation buffer (0.5% BSA, 5% donkey serum and 0.25% Tween-20 in PBS). After washes with PBST, samples were incubated with Alexa-Fluor-conjugated secondary antibodies (Invitrogen, 1:200). Cells were washed and mounted in VectaShield. As a positive control, HUVECs were treated with 50 μ M chloroquine overnight before fixation.

Image acquisition and processing. Stained tissue/cells were analysed at high resolution with a TCS SP8 confocal microscope (Leica). Velocity (Perkin Elmer), Fiji/ImageJ, Photoshop (Adobe) and Adobe Illustrator (Adobe) software were used for image acquisition and processing. For all of the images in which the levels of immunostaining were compared, settings for laser excitation and confocal scanner detection were kept constant between groups.

Quantitative analysis of the retinal and hindbrain vasculature. All quantifications were done on high-resolution confocal images of thin z-sections of the sample using the Velocity (Perkin Elmer) software. In the retina, endothelial coverage, the number of endothelial branchpoints, and the average vessel branch diameter were quantified behind the angiogenic front in a region between an artery and a vein. In the embryonic hindbrain, randomly chosen fields were used to quantify the vascularization in the ventricular zone. All parameters were quantified in a minimum of four vascularized fields per sample. Endothelial coverage was determined by assessing the ratio of the IB4-positive area to the total area of the vascularized field (sized 200 μ m \times 200 μ m), and expressed as a percentage of the area covered by IB4-positive ECs. Average vessel diameter was analysed by assessing the diameter of individual vessel branches in a vascularized field (sized 200 μ m \times 200 μ m), which was used to calculate the mean diameter in each field. The diameter of individual vessel branches was averaged from three measurements taken at the proximal, middle and distal part of the vessel segment. The number of filopodial extensions was quantified at the angiogenic front. The total number of filopodia was normalized to a vessel length of 100 μ m at the angiogenic front, which was defined and measured according to published protocols³⁵. For quantifying vascular outgrowth in the mouse retina, the distance of vessel growth from the centre of the optic nerve to the periphery was measured in each leaflet of a dissected retina, which was used to calculate the mean value for each sample. The number of ERG/IB4- and BrdU/IB4-labelled cells was counted in at least four fields sized 200 μ m \times 200 μ m per sample. Because of the lower incidence of pHH3-positive ECs, the number of pHH3/IB4-double-positive cells was quantified in larger fields (sized 580 μ m \times 580 μ m). For the quantification of the mosaic control (*Pdgfrb-creERT2*; *Rosa26-Foxo1^{+/+}*; *Rosa26-mTmG^{fl/+}*) and *Foxo1^{IEC-CA}* (*Pdgfrb-creERT2*; *Rosa26-Foxo1^{CA/+}*; *Rosa26-mTmG^{fl/+}*) retinas, the GFP/IB4 double-positive area per field was determined and divided by the total IB4-positive area. The percentage of the GFP/IB4 double-positive area per total IB4 area was measured in four fields (400 μ m \times 400 μ m) per sample and used to calculate the mean value. For the quantification of nuclear FOXO1 expression in control and *Foxo1^{IEC-CA}* mice, high-resolution confocal images were taken with a \times 40 objective. The resulting images were analysed with the Bitplane Imaris software. Vessels were first segmented using the Surface module in Imaris. FOXO1 immunofluorescence was then used to set a threshold in the new vascular surface area, in which only CD31-positive nuclei were selected (Surface module). The sum intensity of the nuclear FOXO1 fluorescence was divided by the total vascular area to adjust for differences in vascular density on each image. An average of six images per sample was quantified in three animals per group. All of the images shown are representative of the vascular phenotype observed in samples from at least two distinct litters per group.

Cell culture. Pooled HUVECs were purchased from Lonza and authenticated by marker expression (CD31/CD105 double-positive) and morphology. HUVECs were cultured in endothelial basal medium (EBM; Lonza) supplemented with hydrocortisone (1 μ g ml⁻¹), bovine brain extract (12 μ g ml⁻¹), gentamicin (50 μ g ml⁻¹), amphotericin B (50 ng ml⁻¹), epidermal growth factor (10 ng ml⁻¹) and 10% fetal bovine serum (FBS; Life Technologies). HUVECs were tested negative for mycoplasma and cultured until the fourth passage. The isolation of mouse lung ECs was performed as described³⁶. In brief, adult mice were killed, lungs were removed and incubated with dispase. The homogenate was filtered through a cell strainer, collected by centrifugation, and washed with PBS containing 0.1% BSA (PBSB). The resulting cell suspension was incubated with rat anti-mouse VE-cadherin antibody- (BD Pharmingen, #555289) coated magnetic beads (Dynabeads, Invitrogen, #11035). Next, the beads were washed with PBSB and then resuspended in DMEM/F12 (Invitrogen) supplemented with 20% FCS, endothelial growth factor (Promocell, #C-30140), penicillin and streptomycin. The isolated cells were seeded on gelatin-coated culture dishes and re-purified with the VE-cadherin antibody during the first three passages.

Adenoviral infection. Sub-confluent HUVECs were infected with adenoviruses to overexpress constitutively active human FOXO1-Flag (FOXO1^{CA})³⁷, human

c-MYC-HA³⁸ (Vector Biolabs) and GFP or LacZ as a control. HUVECs (70–80% confluent) were incubated in EBM containing 0.1% BSA for 4 h. Prior to infection, adenoviruses were incubated with an antennapedia-derived peptide (Eurogentec) to facilitate the infection. The mixture was then applied to the HUVECs cultured in EBM containing 0.1% BSA and incubated for 4 h. Thereafter, the cells were washed five times and cultured in EBM with 10% FCS and supplements. The adenoviral infection of murine ECs was performed with adenoviruses encoding for Cre or GFP (Vector Biolabs) as a control.

RNA interference. To silence *FOXO1*, *MYC* or *MXI1* gene expression, HUVECs were transfected with a pool of siRNA duplexes directed against human FOXO1, human c-MYC or human MXI1 (ON-TARGETplus SMARTpool, Dharmacon). A negative control pool of four siRNAs designed and microarray-tested for minimal targeting of human, mouse or rat genes was used as a control (ON-TARGETplus Non-targeting pool, Dharmacon). HUVECs were transfected with 50 nM of the indicated siRNAs using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's recommendations.

Microarray and gene set enrichment analysis. Total RNA quality was verified using the Agilent Bioanalyzer and the 6000 nano kit. RNA was labelled according to the Affymetrix Whole Transcript Sense Target Labelling protocol. Affymetrix GeneChip Human Gene 1.0 ST arrays were hybridized, processed and scanned using the appropriate Affymetrix protocols. Data were analysed using the Affymetrix expression console using the RMA algorithm, statistical analysis was done using DNASTar Arraystar 11. Heat maps were generated using GENE-E, publicly available from the Broad Institute (<http://www.broadinstitute.org/cancer/software/GENE-E/>). For gene set enrichment analysis (GSEA), gene set collections from the Molecular Signatures Database (MSigDB) 4.0 (<http://www.broadinstitute.org/gsea/msigdb/>) were used for the analysis of the endothelial FOXO1 and MYC transcriptomes.

qRT-PCR. RNA was extracted from cells using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. cDNA synthesis was performed on 2 µg of total RNA using the M-MLV reverse transcriptase (Invitrogen). qPCR was performed with TaqMan Gene Expression Master Mix (Applied Biosystems) and TaqMan probes (TaqMan Gene Expression Assays) available from Applied Biosystems. TaqMan Gene Expression Assays used were as follows: human *ACTB* Hs99999903_m1; *CCNB2* Hs00270424_m1; *CCND1* Hs00765553_m1; *CCND2* Hs00153380_m1; *CDK4* Hs00262861_m1; *c-MYC* Hs00153408_m1; *ENO1* Hs00361415_m1; *FASN* Hs01005622_m1; *FBXW7* Hs00217794_m1; *FOXO1* Hs01054576_m1; *LDHA* Hs00855332_g1; *LDHB* Hs00929956_m1; *MXI1* Hs00365651_m1; *PKM2* Hs00987254_m1. Mouse probes were: *Actb* Mm 00607939_s1; *Myc* Mm00487804_m1. All qPCR reactions were run on a StepOnePlus real-time PCR instrument (Applied Biosystems) and data were calculated using the $\Delta\Delta C_t$ method.

Western blot analysis and antibodies. Western blot analyses were performed with precast gradient gels (Bio-Rad) using standard methods. Briefly, HUVECs were lysed in RIPA buffer (150 mM NaCl, 1.0% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS and 50 mM Tris, pH 8.0) supplemented with a protease inhibitor mix (Complete Mini Protease Inhibitor cocktail tablets, Roche) and phenylmethylsulfonyl fluoride. Proteins were separated by SDS-PAGE and blotted onto nitrocellulose membranes (Bio-Rad). Membranes were probed with specific primary antibodies and then with peroxidase-conjugated secondary antibodies. The following antibodies were used: AMPK α (Cell Signaling Technology, #2532, 1:1,000), caspase 3 (Cell Signaling Technology, #9662, 1:1,000), cleaved caspase 3 (Asp175) (Cell Signaling Technology, #9664, 1:1,000), cleaved PARP (Cell Signaling Technology, #5625, 1:1,000), c-MYC (Cell Signaling Technology, #9402, 1:1,000), FBXW7 (Abcam, #12292, 1:500), Flag M2 (Sigma, #F-3165, 1:1,000), FOXO1 (Cell Signaling Technology, #2880, 1:1,000), HA (Covance, clone 16B12, MMS-101P, 1:1,000), LC3A/B (Cell Signaling Technology, #12741, 1:1,000), MXI1 (Santa Cruz, SC-1042, 1:500), P-ACC (Cell Signaling Technology, #3661, 1:1,000), P-AMPK α (Thr 172) (Cell Signaling Technology, #2535, 1:1,000), PARP (Cell Signaling Technology, #9532, 1:1,000), Tubulin (Cell Signaling Technology, #2148, 1:1,000). The bands were visualized by chemiluminescence using an ECL detection kit (Clarity Western ECL Substrate, Bio-Rad) and a ChemiDoc MP Imaging System (Bio-Rad). The gel source data of the western blot analysis is illustrated in Supplementary Fig. 1. Quantification of band intensities by densitometry was carried out using the Image Lab software (Bio-Rad).

Metabolic assays. Extracellular acidification (ECAR) and oxygen consumption (OCR) rates were measured using the Seahorse XFe96 analyser (Seahorse Bioscience) following the manufacturer's protocols. Briefly, ECAR and OCR were measured 4 h after seeding HUVECs (40,000 cells per well) on fibronectin-coated XFe96 microplates. HUVECs were maintained in non-buffered assay medium in a non-CO₂ incubator for 1 h before the assay. The Glycolysis stress test kit (Seahorse Bioscience) was used to monitor the extracellular acidification rate under various conditions. Three baseline recordings were made, followed by sequential injection

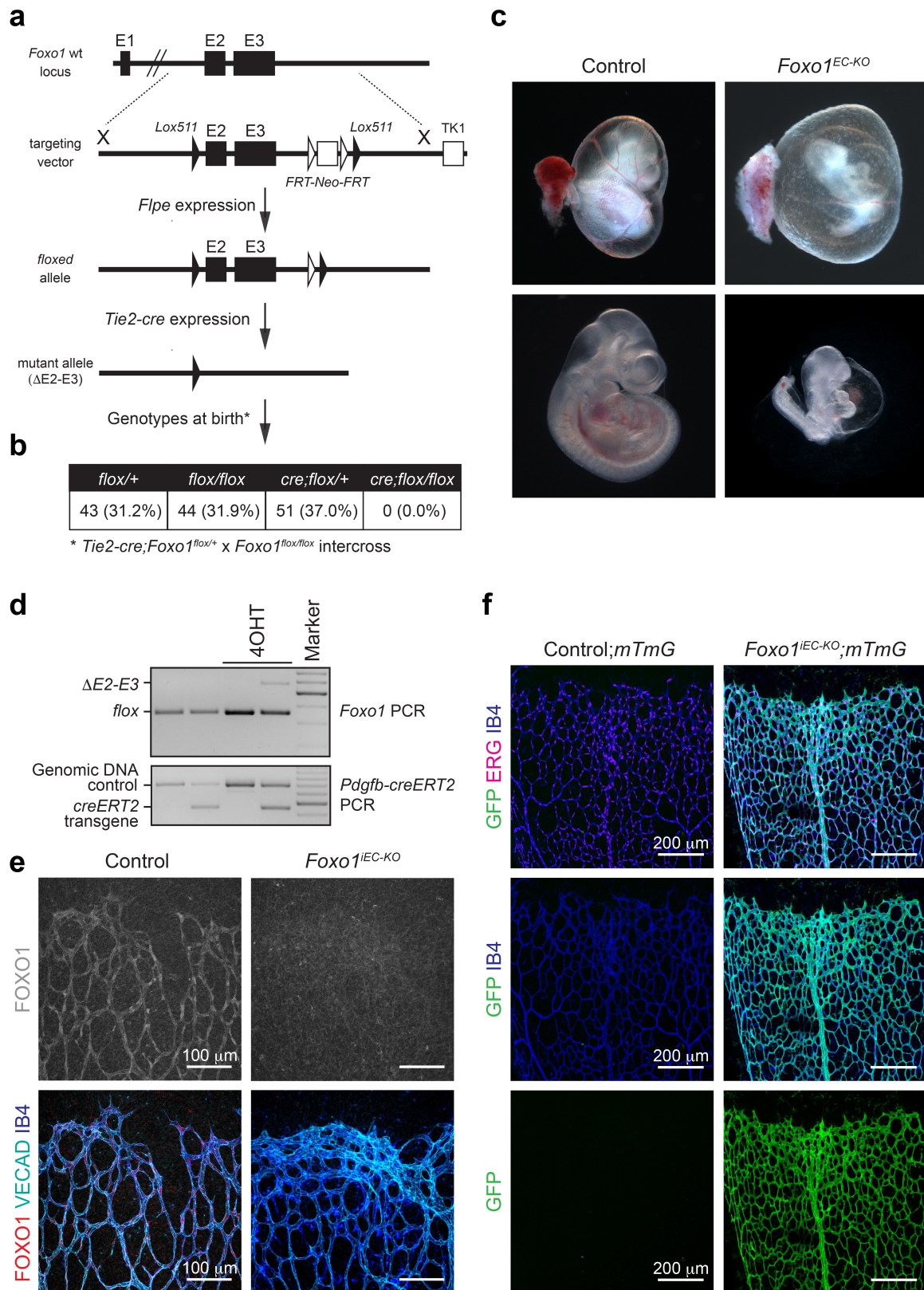
of glucose (10 mM), the mitochondrial/ATP synthase inhibitor oligomycin (3 µM), and the glycolysis inhibitor 2-deoxy-D-glucose (2-DG; 100 mM). The Mito stress test kit was used to assay the mitochondrial respiration rate under basal conditions, in the presence of the ATP synthase inhibitor oligomycin (3 µM), the mitochondrial uncoupler carbonyl cyanide-4-(trifluoromethoxy)phenyl-hydrazone (FCCP; 1 µM), and the respiratory chain inhibitors antimycin A (1.5 µM) and rotenone (3 µM). To measure glycolysis in ECs, HUVECs were incubated for 2 h in growth medium containing 80 µCi mmol⁻¹ [5-³H]-D-glucose (Perkin Elmer). Thereafter, supernatant was transferred into glass vials sealed with rubber stoppers. ³H₂O was captured in hanging wells containing a Whatman paper soaked with H₂O over a period of 48 h at 37 °C to reach saturation⁴. Radioactivity was determined by liquid scintillation counting and normalized to protein content. Lactate concentration in the HUVEC culture media was measured by using a Lactate Assay Kit (Biovision) following the instructions of the manufacturer. Glucose uptake was assessed by analysing the uptake of 2-DG with a Colorimetric Assay (BioVision). ATP was measured from lysates from HUVECs (1 × 10⁶ per ml) with an ATP Bioluminescence Assay Kit CLS II (Roche) according to the instructions of the manufacturer.

Intracellular ROS measurement. Intracellular ROS levels were determined using CM-H₂DCFDA dye (Life technologies). Dye was reconstituted in DMSO (10 mM) and diluted 1:1,000 in PBS containing CaCl₂ and MgCl₂ as working solution. Twenty-four hours after transduction, 1 × 10⁶ cells were incubated in 1 ml working solution for 40 min at 37 °C in the dark. Subsequently the fluorescence of 10,000 living endothelial cells per sample was measured at the BD FACS LSR II flow cytometer. The assays were performed with adenoviruses, which did not co-express fluorescent reporter genes. Data were analysed using BD FACSDiva software (version 8.0.1).

Senescence-associated β -galactosidase staining. To detect senescence-associated β -galactosidase activity in HUVECs, a cellular senescence assay kit (#KAA002, Chemicon) was used according to the manufacturer's instructions. Briefly, cells were fixed in 1 ml fixing solution at room temperature for 15 min. Two millilitres of freshly prepared SA- β -gal detection solution was added and cells were incubated overnight at 37 °C without CO₂ and protected from light. Then the detection solution was removed and cells were washed and mounted in 70% glycerol in PBS. H₂O₂-treated HUVECs were used as a positive control.

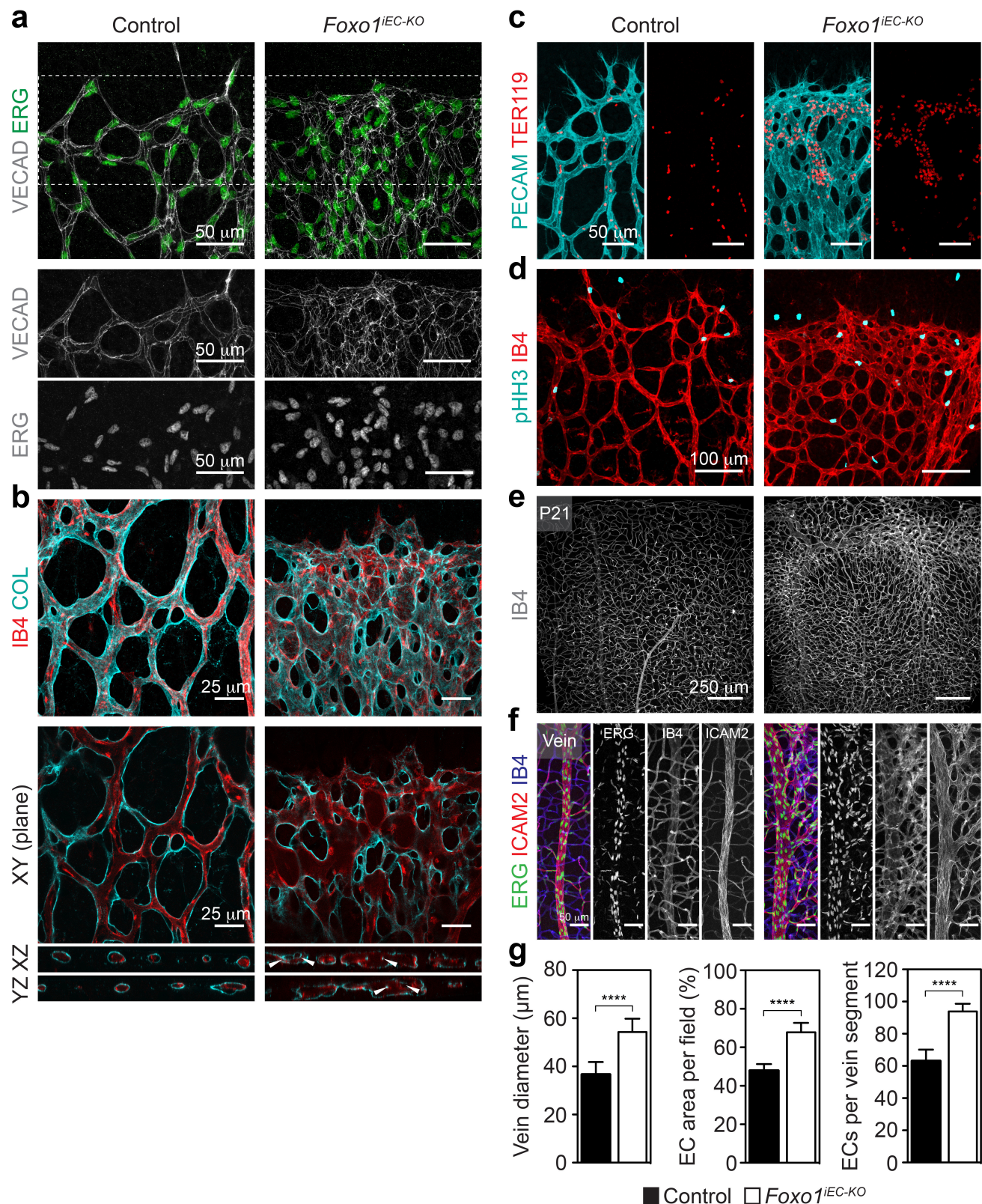
Statistical analysis. Statistical analysis was performed by unpaired, two-tailed Student's *t*-test, or non-parametric one-way ANOVA followed by Bonferroni's multiple comparison test unless mentioned otherwise. For all bar graphs, data are represented as mean \pm s.d. *P* values < 0.05 were considered significant. All calculations were performed using GraphPad Prism software. No randomization or blinding was used and no animals were excluded from the analysis. Sample sizes were selected on the basis of published protocols^{34,35} and previous experiments. Several independent experiments were performed to guarantee reproducibility and robustness of findings.

- Koni, P. A. *et al.* Conditional vascular cell adhesion molecule 1 deletion in mice: impaired lymphocyte migration to bone marrow. *J. Exp. Med.* **193**, 741–754 (2001).
- Claxton, S. *et al.* Efficient, inducible Cre-recombinase activation in vascular endothelium. *Genesis* **46**, 74–80 (2008).
- Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45**, 593–605 (2007).
- Fantini, A., Vieira, J. M., Plein, A., Maden, C. H. & Ruhrberg, C. The embryonic mouse hindbrain as a qualitative and quantitative model for studying the molecular and cellular mechanisms of angiogenesis. *Nature Protocols* **8**, 418–429 (2013).
- Pitulescu, M. E., Schmidt, I., Benedito, R. & Adams, R. H. Inducible gene targeting in the neonatal vasculature and analysis of retinal angiogenesis in mice. *Nature Protocols* **5**, 1518–1534 (2010).
- Guarani, V. *et al.* Acetylation-dependent regulation of endothelial Notch signalling by the SIRT1 deacetylase. *Nature* **473**, 234–238 (2011).
- Ramaswamy, S., Nakamura, N., Sansal, I., Bergeron, L. & Sellers, W. R. A novel mechanism of gene regulation and tumor suppression by the transcription factor FKHR. *Cancer Cell* **2**, 81–91 (2002).
- Hermeking, H. *et al.* Identification of CDK4 as a target of c-MYC. *Proc. Natl Acad. Sci. USA* **97**, 2229–2234 (2000).
- Zeller, K. I., Jegga, A. G., Aronow, B. J., O'Donnell, K. A. & Dang, C. V. An integrated database of genes responsive to the Myc oncogenic transcription factor: identification of direct genomic targets. *Genome Biol.* **4**, R69 (2003).
- Menssen, A. & Hermeking, H. Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes. *Proc. Natl Acad. Sci. USA* **99**, 6274–6279 (2002).
- Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nature Genet.* **40**, 499–507 (2008).
- Bild, A. H. *et al.* Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* **439**, 353–357 (2006).



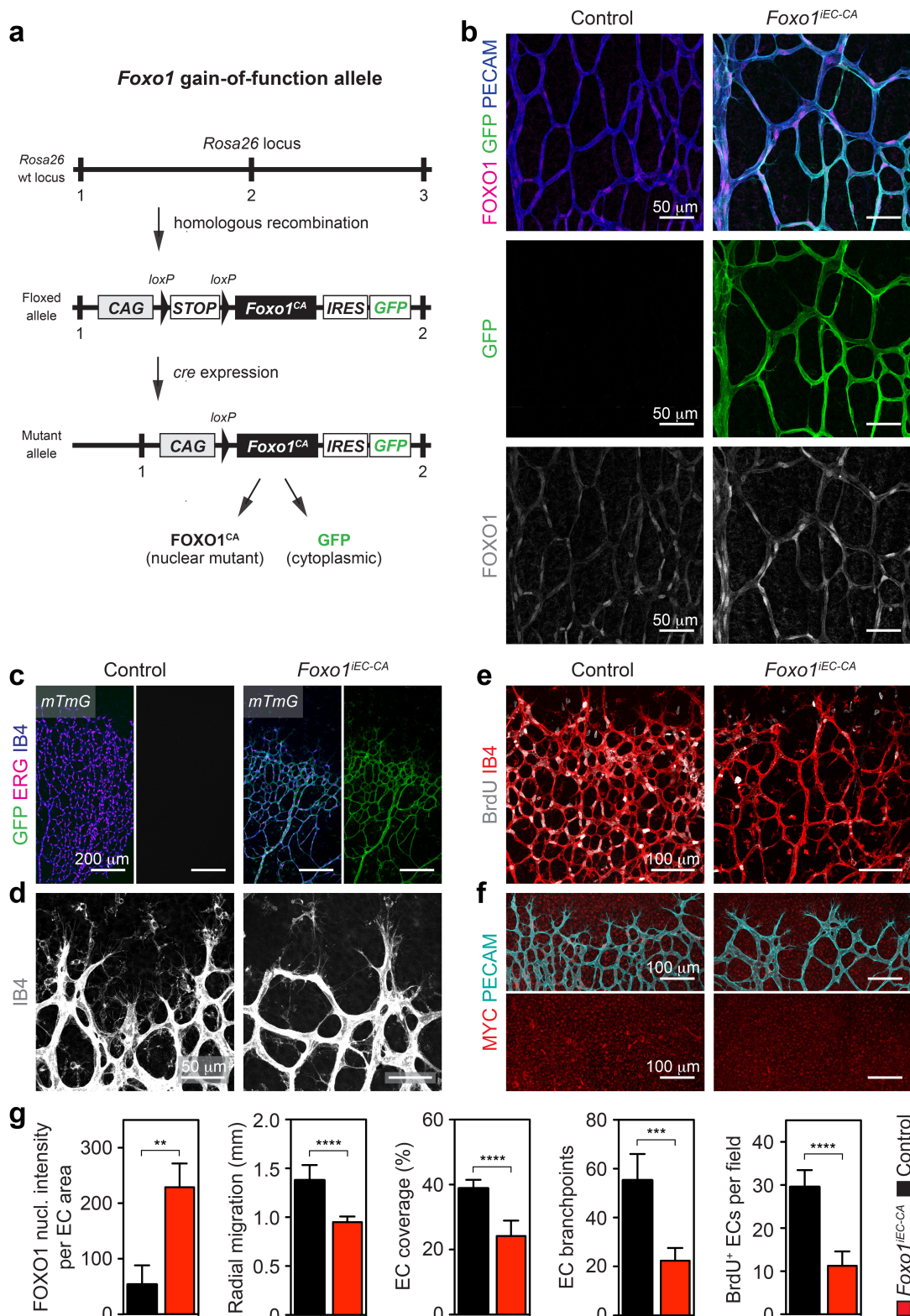
Extended Data Figure 1 | Constitutive and inducible deletion of *Foxo1* in ECs of mice. **a**, Strategy to generate a conditional *Foxo1* mutant allele in which exons 2 and 3 are flanked by *lox* sites. The structures of the genomic locus, the targeting vector, and the targeted allele are shown. *FRT-Neo-FRT*, neomycin resistance cassette flanked by *FRT* sites. TK1, thymidine kinase. **b**, Table of viable offspring from *Tie2-cre*; *Foxo1*^{*flox*/+} (male) and *Foxo1*^{*flox*/*flox*} (female) intercrosses. **c**, Control (*Foxo1*^{*flox*/+}) and *Foxo1*^{*IEC-KO*} mutants (*Tie2-cre*; *Foxo1*^{*flox*/*flox*}) at E10.5. **d**, PCR of genomic DNA from P5 control (*Foxo1*^{*flox*/+}, lanes 1 and 3) and *Foxo1*^{*IEC-KO*} (*Pdgfb-creERT2*; *Foxo1*^{*flox*/*flox*},

lanes 2 and 4) pups untreated (lanes 1 and 2) or treated (lanes 3 and 4) with 4-OHT. Recombination of the floxed *Foxo1* allele (Δ) occurs only in 4-OHT-injected animals that are *Pdgfb-creERT2*-positive. **e**, Immunofluorescence staining for FOXO1, VE-cadherin (VECAD) and isolectin-B4 (IB4) in a P5 mouse retina of 4-OHT-injected control and *Foxo1*^{*IEC-KO*} mice. **f**, Confocal images of *mTmG*⁺ control- and *Foxo1*^{*IEC-KO*} mice that were injected with 4-OHT from P1 to P4 and analysed for GFP, ERG and IB4 expression.



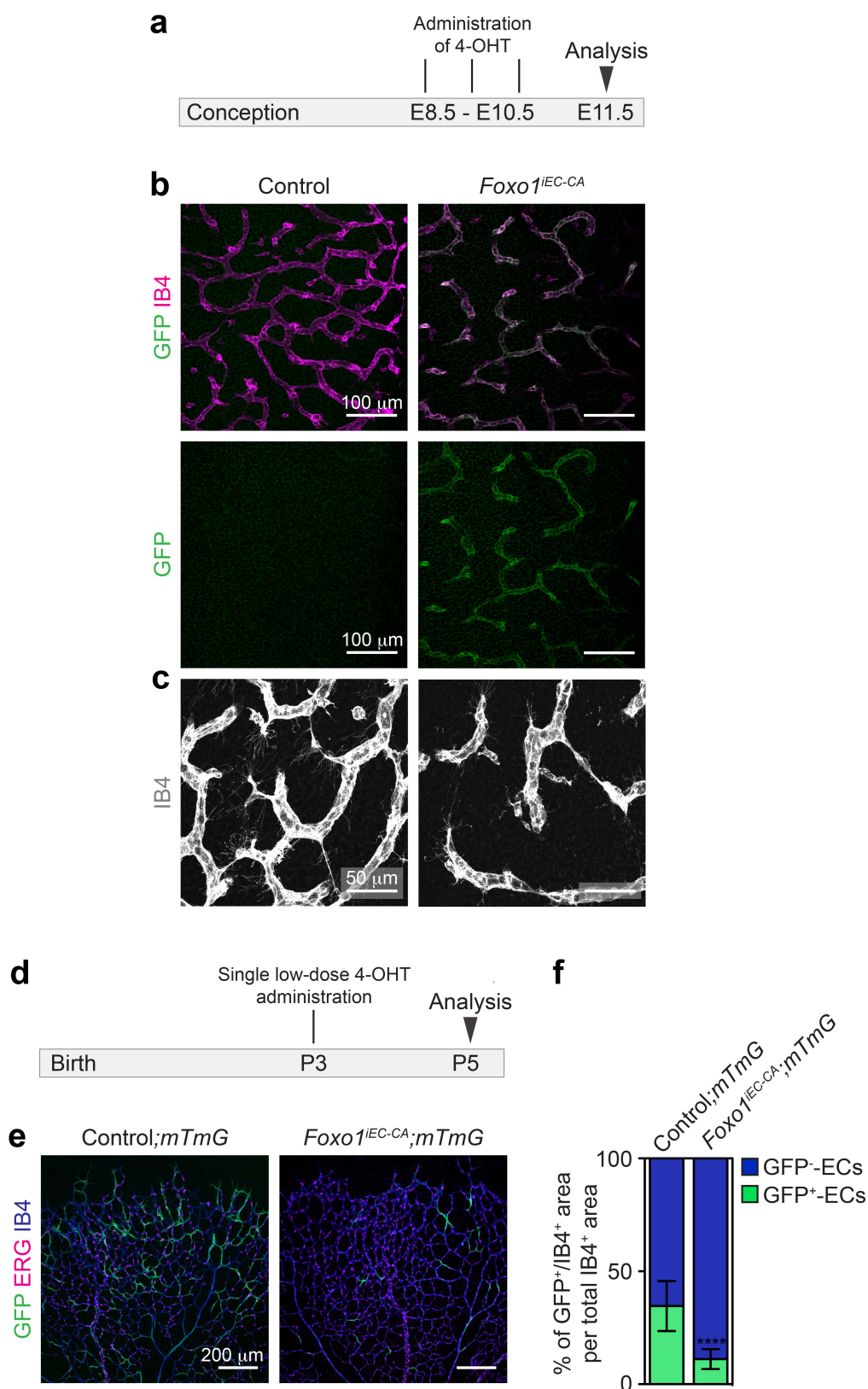
Extended Data Figure 2 | Endothelial FOXO1 deficiency leads to abnormal vessel size and shape. **a**, Immunostaining for VECAD and ERG in *Foxo1*^{IEC-KO} and control retinas. The bottom panels show the isolated VECAD and ERG signals of the inset. **b**, Confocal images showing maximum intensity projections and X-Y, X-Z, and Y-Z planes of a thick stack of IB4 and collagen IV (COL) stained P5 retinas. *Foxo1*^{IEC-KO} mice develop enlarged vessels with abnormal lumen organization. White arrowheads point to areas with multiple vessel layers and intraluminal collagen strands. **c**, Images of IB4- (cyan) and TER119- (red) stained P5 retinas of control and *Foxo1*^{IEC-KO} mice. Note that aggregates of TER119⁺

red blood cells form in *Foxo1*^{IEC-KO} but not in control mice. **d**, Phosphohistone H3 (pH3) and IB4 immunostaining of P5 *Foxo1*^{IEC-KO} and control mice. **e**, Images of IB4-stained retinas at P21 showing an increased vessel density in *Foxo1*^{IEC-KO} mice (same samples as in Fig. 1h). **f**, Higher magnification images of ERG-, ICAM2- and IB4-stained retinas at P21 showing increased numbers of ECs in the perivenous plexus of *Foxo1*^{IEC-KO} mice. **g**, Bar graphs showing the mean endothelial area ($n \geq 8$), mean diameter of central vein ($n \geq 8$), and number of ERG/IB4⁺ cells ($n \geq 4$) in P21 retinas of *Foxo1*^{IEC-KO} and control mice. Data represent mean \pm s.d. Two-tailed unpaired *t*-test. **** $P < 0.0001$.



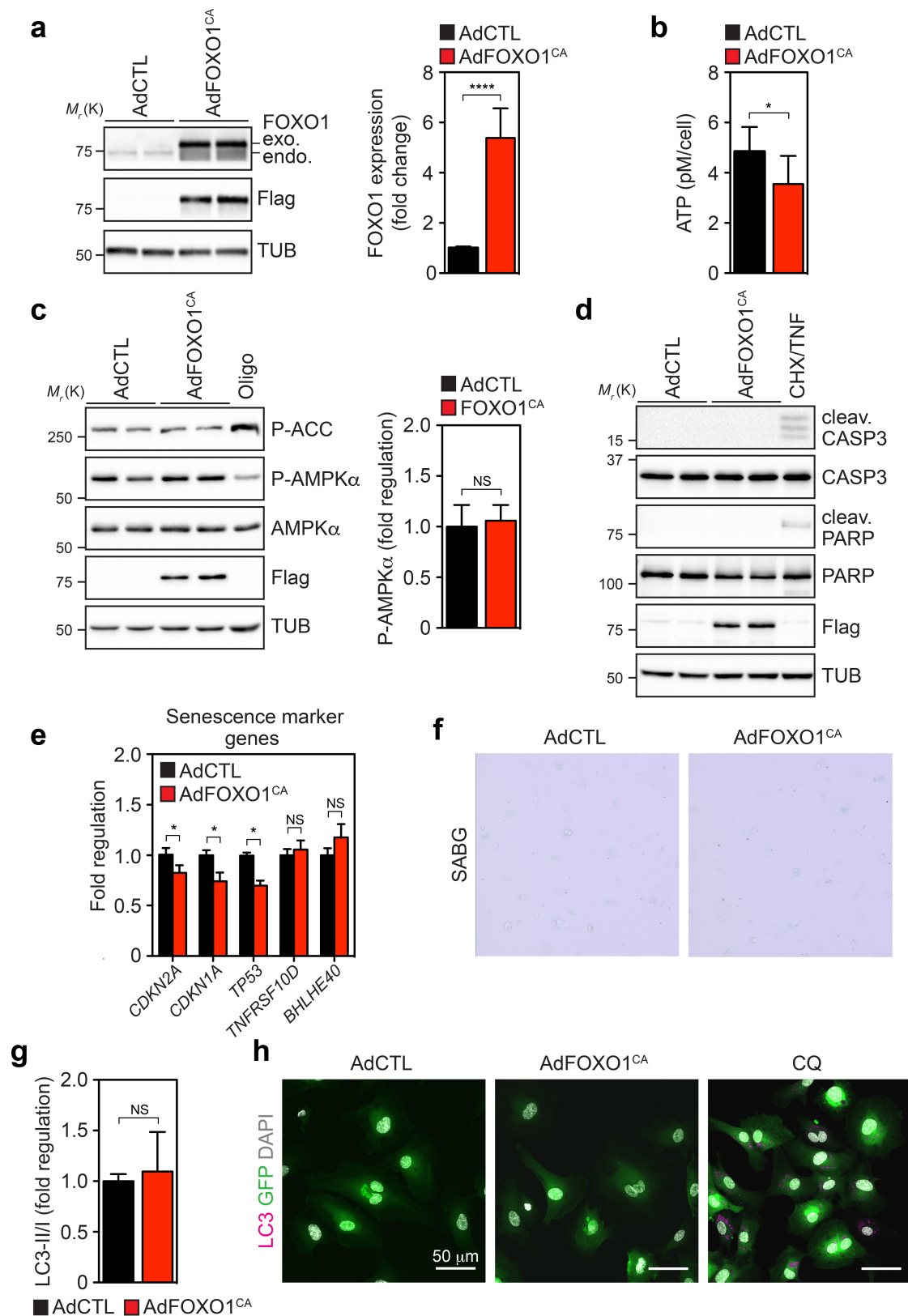
Extended Data Figure 3 | Inducible overexpression of a constitutively active FOXO1 mutant in ECs of mice. **a**, A cassette containing the CAG promoter, a floxed STOP sequence, a cDNA encoding for *Foxo1*^{CA}, and IRES-GFP was inserted into the *Rosa26* locus. A schematic representation of the wild-type *Rosa26* locus, the floxed allele, and the recombined allele after *cre* expression is shown. **b**, Immunofluorescence staining for FOXO1, GFP and PECAM in P5 *Foxo1*^{IEC-CA} and control mice. **c**, Confocal images of *mTmG*⁺ control and *Foxo1*^{IEC-CA} mice that were injected with 4-OHT from P1 to P4 and analysed for GFP, ERG and IB4 expression. The right half of both images shows the GFP signal alone. **d**, High-magnification images of IB4-stained retinal vessels at the angiogenic front in control

and *Foxo1*^{IEC-CA} pups. **e**, BrdU and IB4 labelling of whole-mount P5 retinas reveals reduced endothelial proliferation in *Foxo1*^{IEC-CA} animals. **f**, Confocal images showing MYC and PECAM immunostaining in P5 retinas of control and *Foxo1*^{IEC-CA} mice. The lower half of both images shows the MYC signal alone. **g**, Quantification of FOXO1 nuclear staining intensity in ECs ($n = 3$), radial migration ($n = 10$), endothelial coverage ($n = 10$), branch points ($n = 10$), and endothelial BrdU incorporation ($n \geq 6$) in P5 retinas of control and *Foxo1*^{IEC-CA} mutant mice. Data represent mean \pm s.d. Two-tailed unpaired *t*-test. ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.



Extended Data Figure 4 | FOXO1 restricts EC propagation and vascular growth in a cell-autonomous manner. **a**, Timeline for the analysis of angiogenesis in the embryonic hindbrain. Plug-positive female mice were injected with 4-OHT from E8.5 to 10.5 and embryos harvested on E11.5 for hindbrain dissection. **b**, Confocal images of E11.5 control and *Foxo1^{IEC-CA}* hindbrains stained with IB4 and GFP. **c**, High-magnification images of IB4-stained blood vessels in the ventricular zone of control and *Foxo1^{IEC-CA}* mice. **d**, Timeline for the analysis of control and *Foxo1^{IEC-CA}*

low-degree chimaeras that heterozygously co-express the *mTmG* Cre reporter. Control and *Foxo1^{IEC-KO}* mice were injected with a single low dose of 4-OHT at P3 and retinas analysed at P5. **e**, **f**, Confocal images (**e**) and quantification (**f**) of control;*mTmG* and *Foxo1^{IEC-CA}*;*mTmG* retinas after low-dose 4-OHT treatment at P3 (*n* = 9). Samples were labelled for GFP, ERG and IB4. Data represent mean \pm s.d., two-tailed unpaired *t*-test. *****P* < 0.0001.

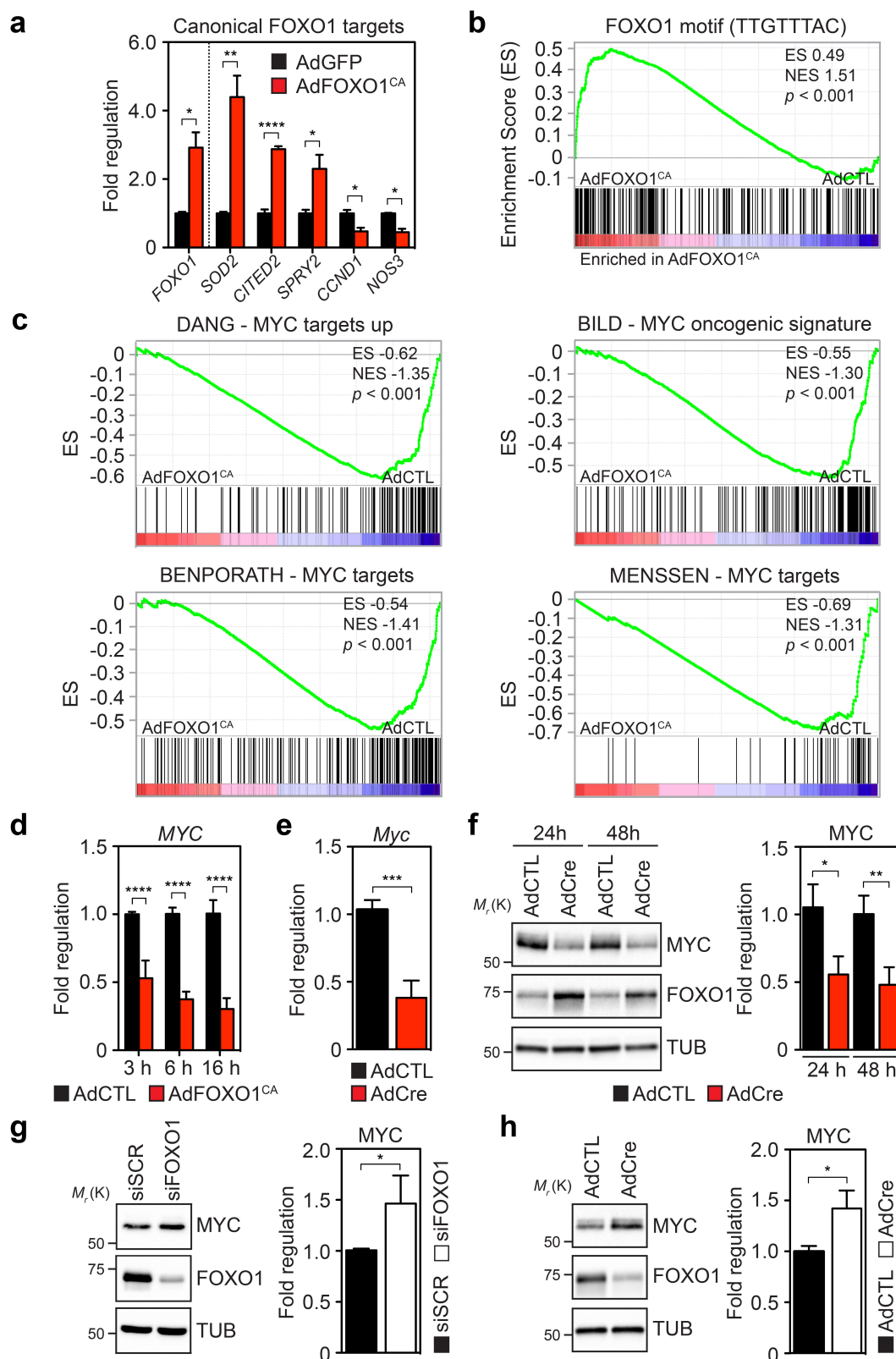


Extended Data Figure 5 | See next page for caption.

Extended Data Figure 5 | Forced expression of FOXO1 does not induce apoptosis, senescence, autophagy or energy distress in cultured ECs.

a, Immunoblot analysis and quantification of FOXO1 protein levels in AdCTL and AdFOXO1^{CA}-Flag transduced HUVECs ($n = 20$). **b**, ATP levels in ECs 24 h after transduction with AdCTL or AdFOXO1^{CA} ($n = 7$). **c**, Western blot images and quantification of AdCTL- or AdFOXO1^{CA}-Flag-transduced HUVECs showing that FOXO1^{CA} does not alter the phosphorylation of AMPK α (Thr 172) or of its substrate ACC (Ser 79). Oligomycin (Oligo), positive control. TUB, tubulin. $n = 10$. **d**, Western blotting of AdCTL- or AdFOXO1^{CA}-Flag-transduced HUVECs illustrating that overexpression of FOXO1^{CA} does not induce apoptotic cell death. Cleaved caspase3 (CASP3) and PARP served as markers of apoptosis. Cycloheximide (CHX) and TNF- α (TNF) costimulation,

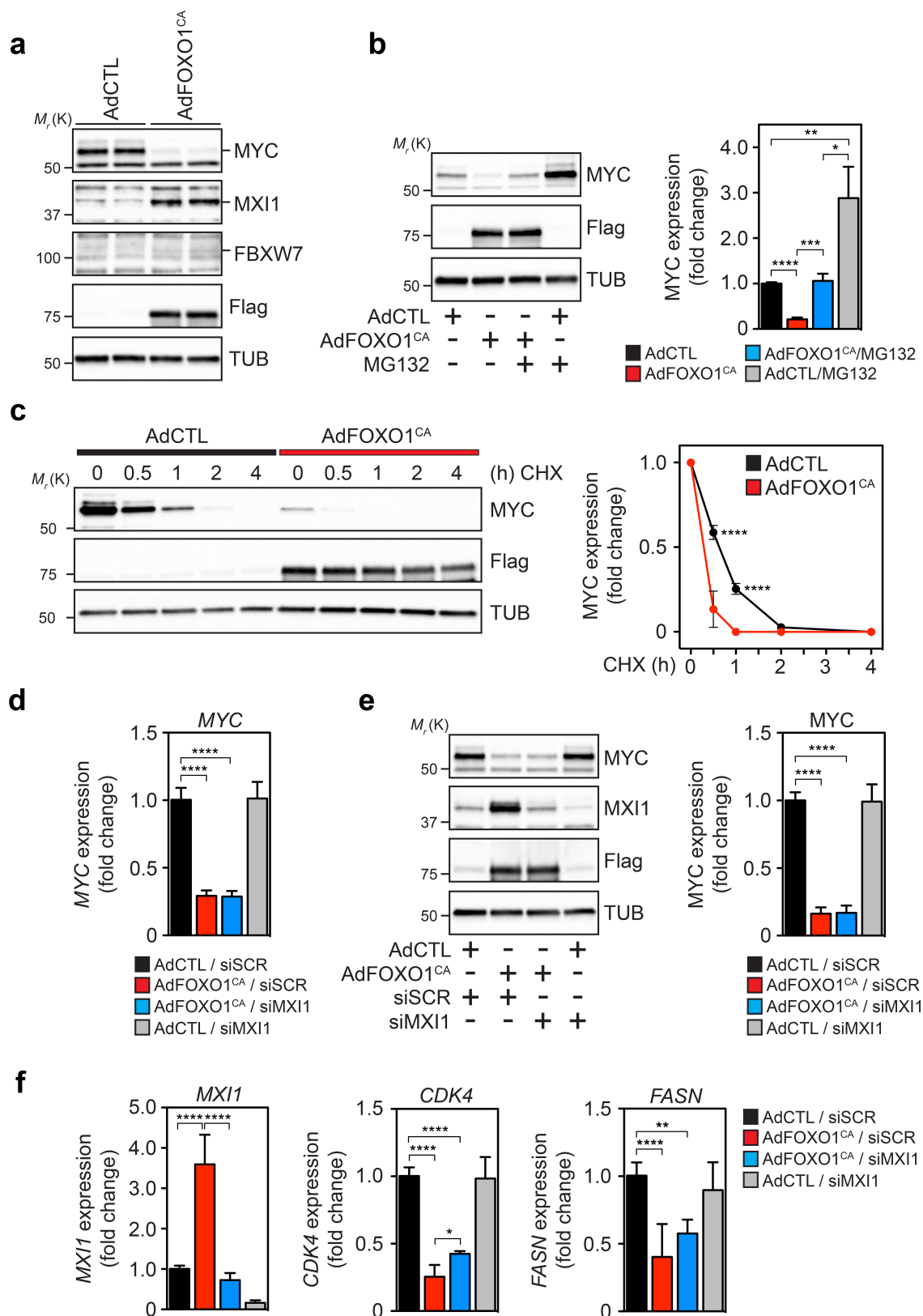
positive control. **e**, Analysis of senescence-associated genes by microarray demonstrating that senescence markers were not significantly changed or even downregulated in FOXO1^{CA}-overexpressing ECs. $n = 3$. **f**, Images of β -galactosidase stainings in AdCTL- and AdFOXO1^{CA}-transduced HUVECs showing no increase in senescence-associated β -galactosidase activity (SABG). **g**, Densitometric quantification of the LC3-II to LC3-I ratio in AdCTL- or AdFOXO1^{CA}-transduced HUVECs ($n = 10$). **h**, Immunofluorescence analysis of AdCTL- and AdFOXO1^{CA}-transduced HUVECs (both coexpressing GFP) using LC3 and GFP antibodies. Chloroquine (CQ), positive control. DAPI, endothelial nuclei. Data in **a–c**, **e** and **g** represent mean \pm s.d. Two-tailed unpaired t -test. * $P < 0.05$; *** $P < 0.0001$; NS, not significant.



Extended Data Figure 6 | FOXO1 represses MYC signalling in ECs.

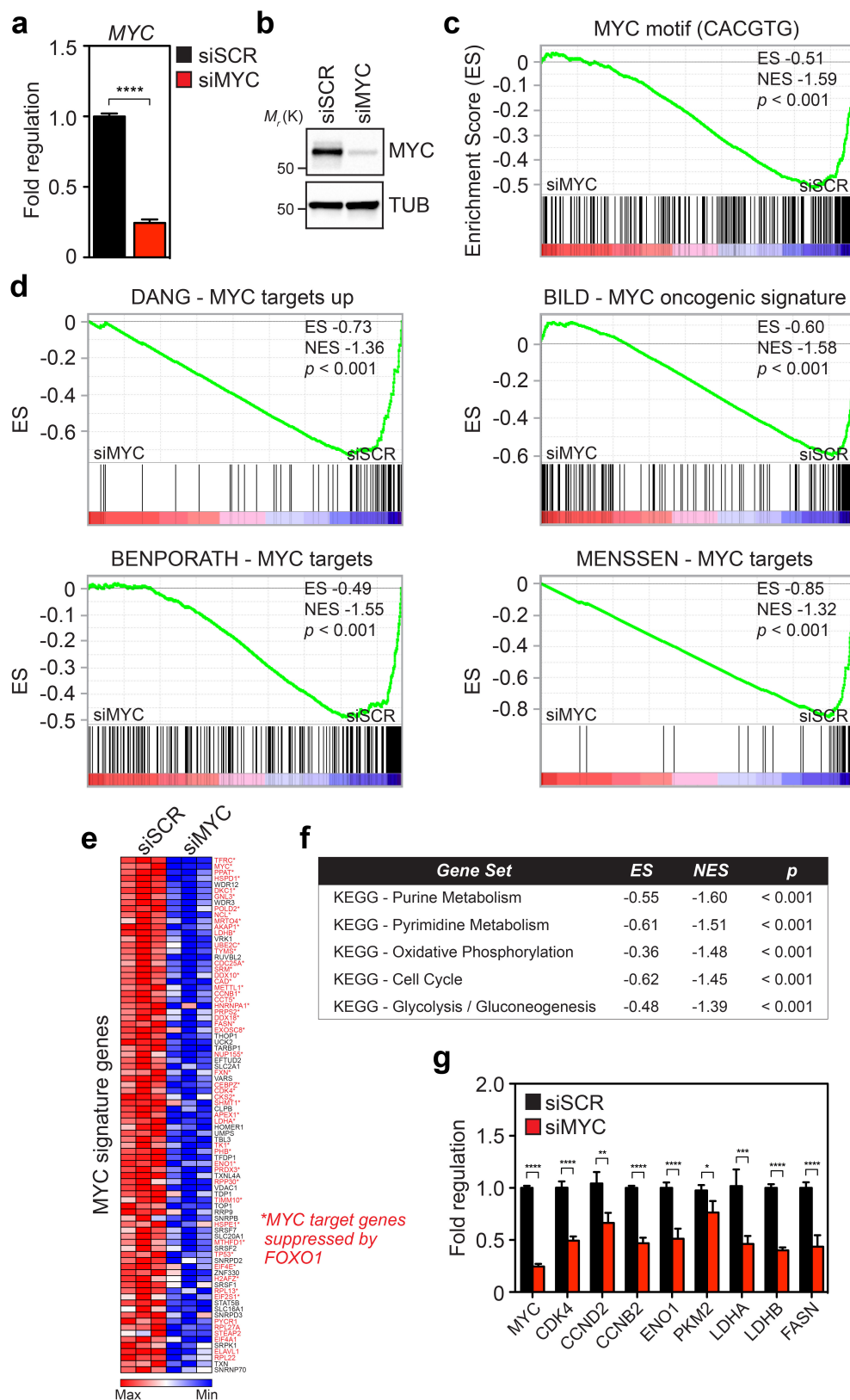
a, Microarray expression analysis of FOXO1 and of canonical FOXO target genes in AdFOXO1^{CA}- and AdCTL-expressing HUVECs 16 h after transduction ($n = 3$). **b**, GSEA of the FOXO1 DNA-binding element (TTGTTTAC) gene set in AdFOXO1^{CA}- or AdCTL-transduced ECs. ES, enrichment score; NES, normalized enrichment score. **c**, GSEA of MYC gene signatures^{39–42} showing the downregulation of MYC target genes in FOXO1^{CA}-expressing HUVECs. **d**, qPCR expression analysis of MYC at 3, 6 and 16 h in AdCTL and AdFOXO1^{CA}-transduced HUVECs ($n \geq 4$). **e**, qPCR analysis ($n = 4$) of Myc mRNA levels in ECs isolated from Foxo1^{CA} mice 24 h

after transduction with a control or Cre (AdCre) adenovirus. **f**, Immunoblot analysis of MYC in ECs isolated from Foxo1^{CA} mice following transduction with AdCTL or AdCre ($n = 3$). Cre-mediated recombination gave rise to a 2.8 ± 0.3 -fold increase in FOXO1 protein expression. **g**, Expression analysis of MYC in HUVECs by western blotting after RNA interference (RNAi)-mediated knockdown of FOXO1 (siFOXO1). siSCR, scrambled control ($n = 3$). **h**, MYC protein expression in ECs isolated from Foxo1^{fl/fl} mice 24 h after transduction with an AdCTL or AdCre-encoding adenovirus ($n = 3$). **a**, **d**–**h**, Data represent mean \pm s.d., two-tailed unpaired t -test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.



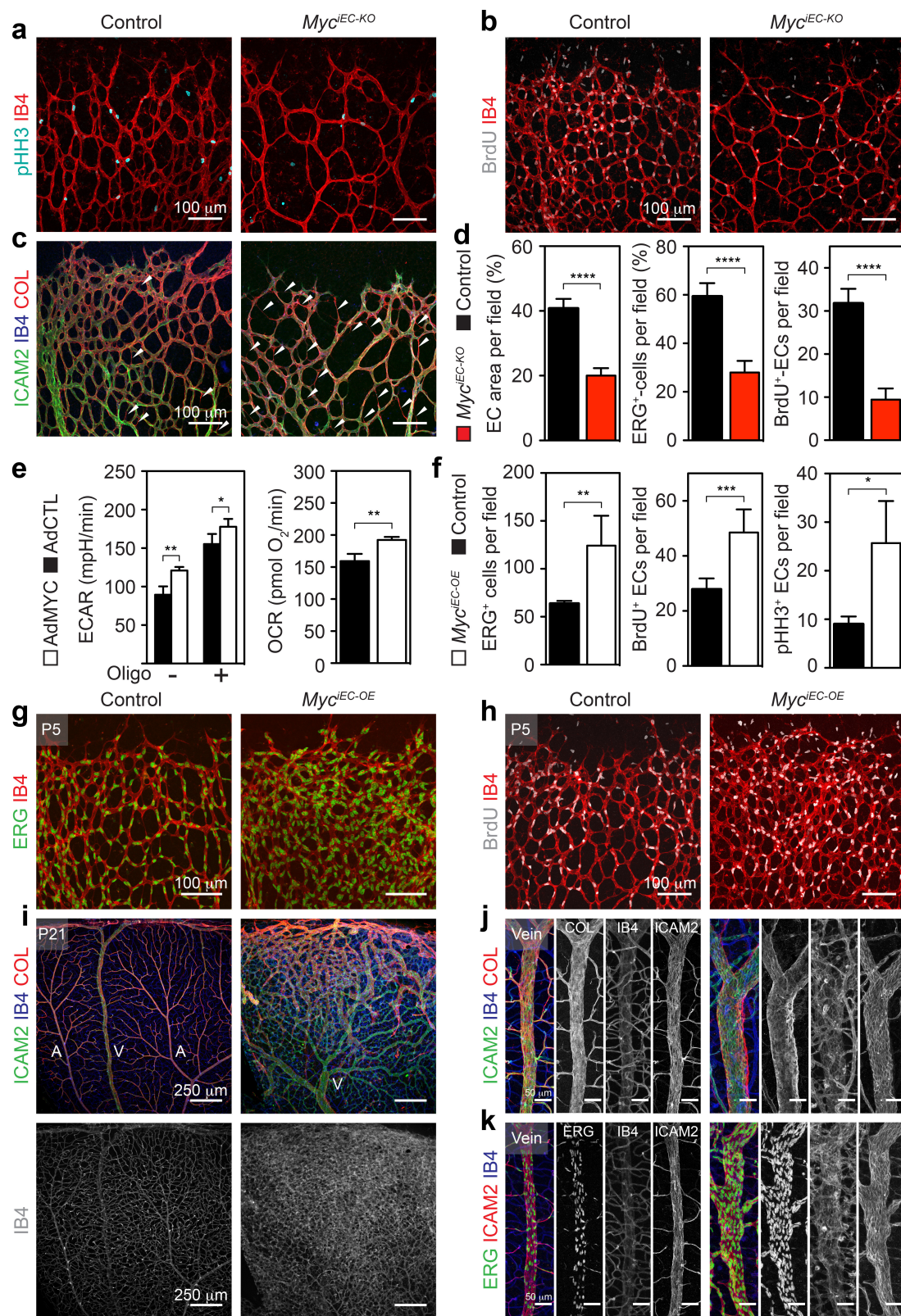
Extended Data Figure 7 | FOXO1 interferes with MYC signalling at different levels. a, Western blot analysis of MYC, MXI1 and FBXW7 in AdCTL- or AdFOXO1^{CA}-Flag-transduced HUVECs. **b**, Immunoblot analysis and quantification of MYC protein levels in AdCTL and AdFOXO1^{CA}-Flag transduced HUVECs that were co-treated with the proteasomal inhibitor MG132 ($n = 3$). **c**, Analysis of MYC protein half-life in AdCTL- or AdFOXO1^{CA}-Flag-transduced HUVECs. The day after transduction, HUVECs were treated with cycloheximide (CHX) and incubated for the times indicated. Data represent mean \pm s.d. Two-way

ANOVA with Bonferroni's multiple comparison post-hoc test. **d**, **e**, qPCR (**d**) and immunoblot analysis (**e**) of MYC levels in control (siSCR) or MXI1 (siMXI1) siRNA-transfected HUVECs that were also transduced with AdCTL or AdFOXO1^{CA}-Flag ($n \geq 5$). **f**, qPCR analysis of MYC target genes in siSCR or siMXI1-transfected HUVECs that were cotransduced with AdCTL or AdFOXO1^{CA} ($n \geq 3$). Data represent mean \pm s.d. One-way ANOVA with Bonferroni's multiple comparison post-hoc test was performed in **b**, **d**, **e** and **f**. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.



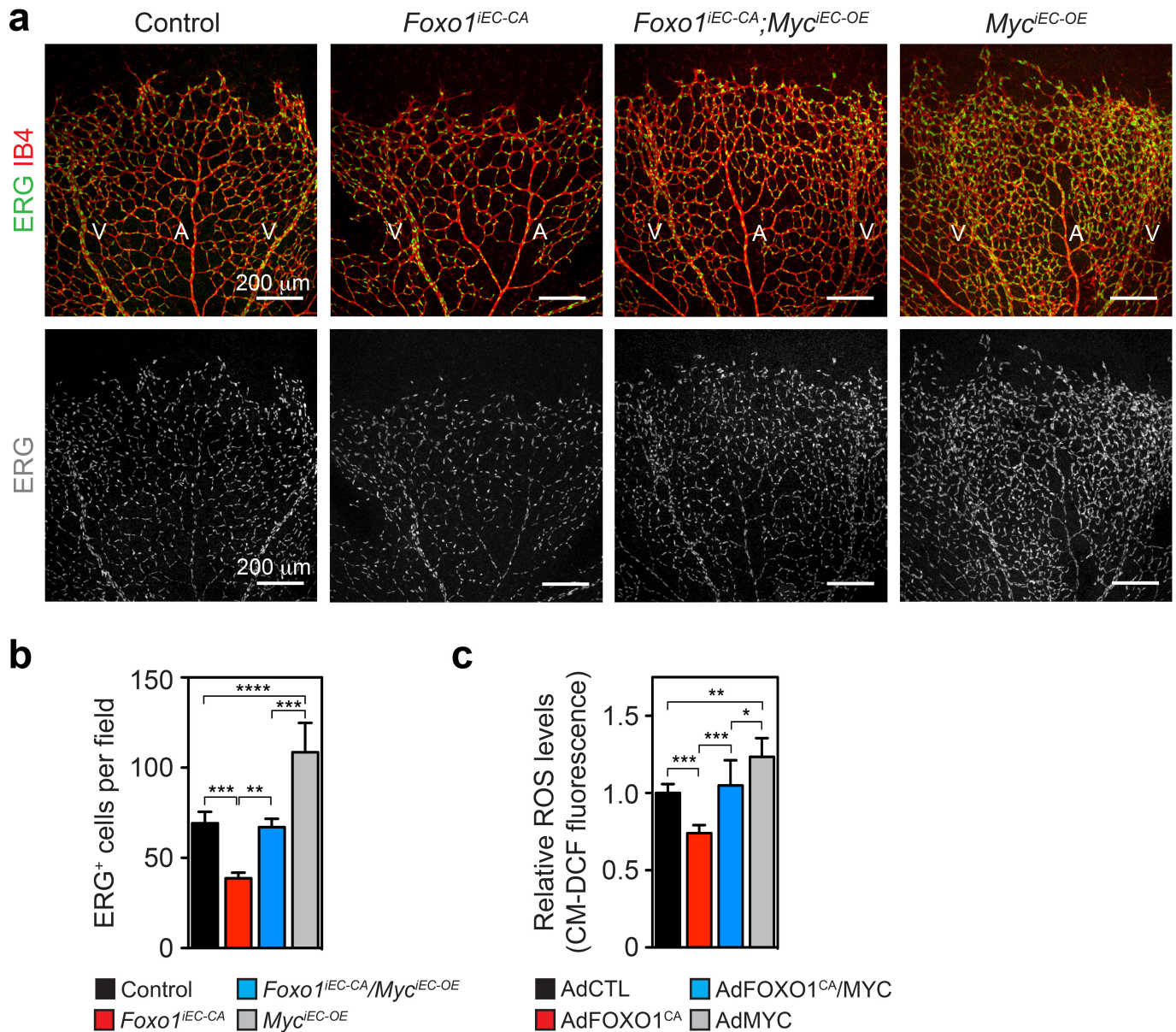
Extended Data Figure 8 | MYC regulates genes involved in cell metabolism and growth in ECs. **a, b,** Analysis of MYC expression by qPCR (**a**) and immunoblot (**b**) in scrambled (siSCR) and MYC (siMYC) siRNA-treated HUVECs 24 h after transfection ($n = 7$). **c,** GSEA of the MYC (CACGTG) DNA-binding element gene set in siSCR- or siMYC-transfected HUVECs. **d,** GSEA of MYC gene signatures^{39–42} showing the downregulation of MYC target genes in MYC-depleted HUVECs. **e,** Heat map of downregulated MYC signature genes in MYC-silenced

HUVECs ($n = 3$). Genes highlighted in red indicate genes that are also suppressed by FOXO1^{CA} overexpression. **f,** Table of KEGG gene sets enriched among genes downregulated in the MYC siRNA-transfected ECs. **g,** Expression analysis of FOXO1-regulated MYC target genes by qPCR in MYC-silenced HUVECs ($n \geq 4$). **a, g,** Data represent mean \pm s.d., two-tailed unpaired t -test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.



Extended Data Figure 9 | MYC is a critical driver of endothelial proliferation, growth and metabolism. **a, b**, IB4 and pHH3 (**a**) or BrdU (**b**) labelling of P5 retinas reveals reduced endothelial proliferation in *Myc^{IEC-KO}* mice. **c**, ICAM2, IB4 and COL staining of retinas at P5 showing an increased number of empty (COL⁺, IB4⁻) sleeves (white arrows) in the plexus of *Myc^{IEC-KO}* mutants. **d**, Quantitative analysis of the indicated vascular parameters in P5 retinas of control and *Myc^{IEC-KO}* mice ($n \geq 8$). **e**, ECAR ($n = 4$) and OCR ($n = 4$) in AdMYC-transduced HUVECs showing a heightened metabolic activity in MYC-overexpressing ECs (6.8 ± 1.4 -fold MYC overexpression). **f**, *Pdgfr-creERT2*-mediated

overexpression of MYC (2.4 ± 0.8 -fold MYC overexpression) enhances vascular growth as indicated by the parameters assessed at P5 ($n \geq 6$). **g**, ERG and IB4 labelling of P5 retinas showing an increase in cellularity in vessels of *Myc^{IEC-OE}* mice. **h**, Enhanced EC proliferation in *Myc^{IEC-OE}* mice as revealed by BrdU and IB4 costaining. **i, j**, Overview (**i**) and higher magnification images (**j**) of ICAM2-, IB4- and COL-stained retinas at P21 showing aberrant vascular growth and venous enlargement in *Myc^{IEC-OE}* mice. **k**, Increased endothelial cellularity in veins of *Myc^{IEC-OE}* mice at P21. **d-f**, Data represent mean \pm s.d., two-tailed unpaired *t*-test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.



Extended Data Figure 10 | Restoration of MYC signalling in FOXO1^{CA}-overexpressing endothelium normalizes vascular growth. a, b, Confocal images (a) and quantification (b) of ERG- and IB4-stained P5 retinas in control, *Foxo1*^{IEC-CA}, *Myc*^{IEC-OE} and *Foxo1*^{IEC-CA}/*Myc*^{IEC-OE} mice (same samples as in Fig. 4h) showing that EC numbers are normalized in the *Foxo1*^{IEC-CA}/*Myc*^{IEC-OE} double mutants ($n \geq 3$). c, Relative ROS levels in

AdCTL-, AdFOXO1^{CA}-, AdFOXO1^{CA}/AdMYC- and AdMYC-transduced HUVECs showing that ROS levels increase again in FOXO1^{CA}/MYC co-expressing ECs ($n \geq 6$). **b, c**, Data represent mean \pm s.d., one-way ANOVA with Bonferroni's multiple comparison post-hoc test. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

Tuft-cell-derived IL-25 regulates an intestinal ILC2-epithelial response circuit

Jakob von Moltke¹, Ming Ji^{1,2}, Hong-Erh Liang¹ & Richard M. Locksley^{1,2,3}

Parasitic helminths and allergens induce a type 2 immune response leading to profound changes in tissue physiology, including hyperplasia of mucus-secreting goblet cells¹ and smooth muscle hypercontractility². This response, known as 'weep and sweep', requires interleukin (IL)-13 production by tissue-resident group 2 innate lymphoid cells (ILC2s) and recruited type 2 helper T cells (T_H2 cells)³. Experiments in mice and humans have demonstrated requirements for the epithelial cytokines IL-33, thymic stromal lymphopoietin (TSLP) and IL-25 in the activation of ILC2s^{4–11}, but the sources and regulation of these signals remain poorly defined. In the small intestine, the epithelium consists of at least five distinct cellular lineages¹², including the tuft cell, whose function is unclear. Here we show that tuft cells constitutively express IL-25 to sustain ILC2 homeostasis in the resting lamina propria in mice.

After helminth infection, tuft-cell-derived IL-25 further activates ILC2s to secrete IL-13, which acts on epithelial crypt progenitors to promote differentiation of tuft and goblet cells, leading to increased frequencies of both. Tuft cells, ILC2s and epithelial progenitors therefore comprise a response circuit that mediates epithelial remodelling associated with type 2 immunity in the small intestine, and perhaps at other mucosal barriers populated by these cells.

To study the source and regulation of IL-25 *in vivo*, we generated a knock-in mouse termed Flare25 (flox and reporter of *Il25*; *Il25*^{F25/F25}) that expresses tandem-dimer red fluorescent protein (RFP) from the *Il25* locus and enables conditional deletion of IL-25 activity (Extended Data Fig. 1a). Immunohistochemistry and flow cytometry revealed RFP only in rare epithelial (epithelial cell adhesion molecule (EPCAM)⁺) cells throughout the digestive tract (Fig. 1a, b and

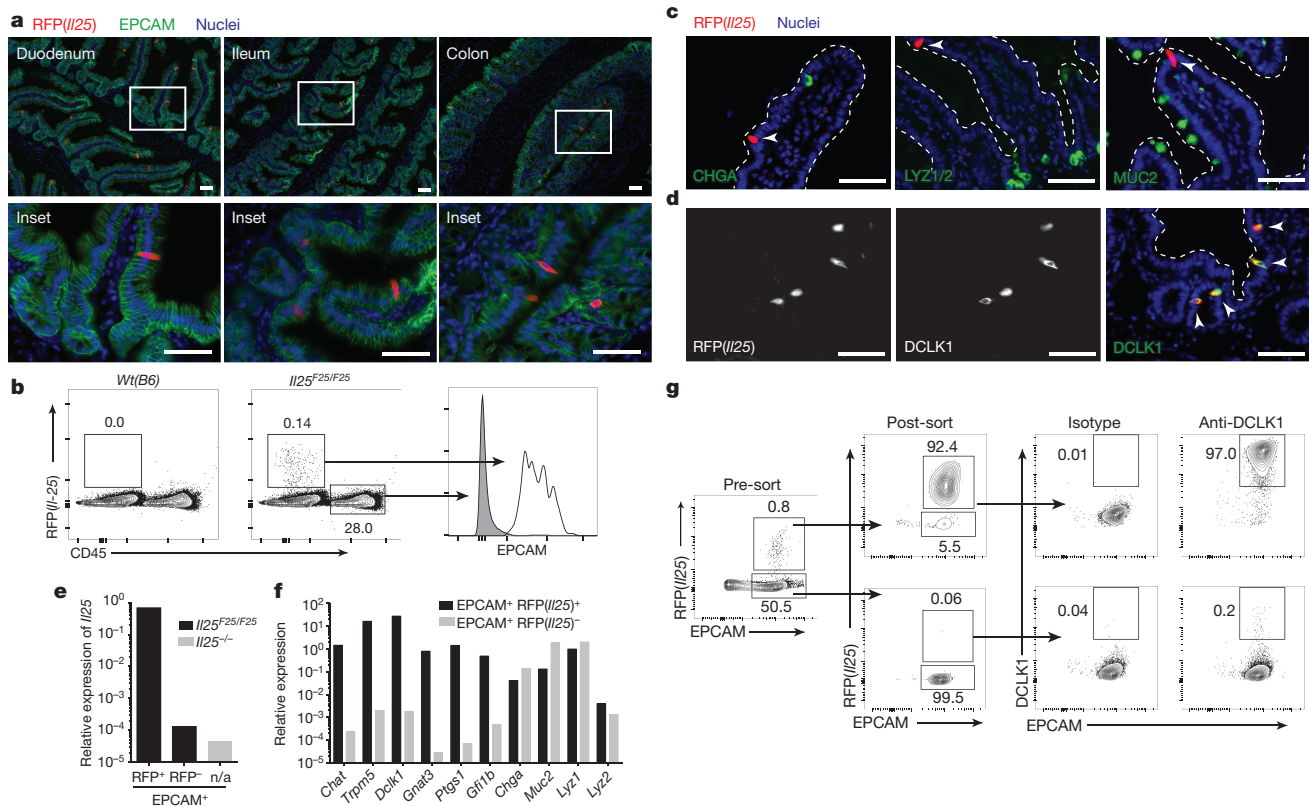


Figure 1 | Intestinal tuft cells constitutively express *Il25*. **a**, Indicated tissues from *Il25*^{F25/F25} mice stained for RFP (red), EPCAM (green) and 4',6-diamidino-2-phenylindole (DAPI; blue). **b**, Flow cytometry of digested jejunum. **c**, **d**, Jejunum from *Il25*^{F25/F25} mice stained as indicated. Dotted lines outline villi. Arrowheads indicate RFP⁺ cells. **e**, **f**, Quantitative polymerase chain reaction with reverse transcription (RT-PCR) on cells

sorted from small intestines of *Il25*^{-/-} (**e**) and *Il25*^{F25/F25} (**e**, **f**) mice. n/a, not applicable. **g**, Flow cytometry of cells sorted from small intestines of *Il25*^{F25/F25} mice and stained with anti-DCLK1. Scale bars, 50 μ m. All data are biological replicates. Data are representative of two (**b–d**, **g**), or at least three (**a**, **e**, **f**) experiments. In **a**, $n > 5$; in **b–d**, $n = 2$; in **e**, $n = 3$.

¹Department of Medicine, University of California San Francisco, San Francisco, California 94143-0795, USA. ²Howard Hughes Medical Institute, University of California San Francisco, San Francisco, California 94143-0795, USA. ³Department of Microbiology & Immunology, University of California San Francisco, San Francisco, California 94143-0795, USA.

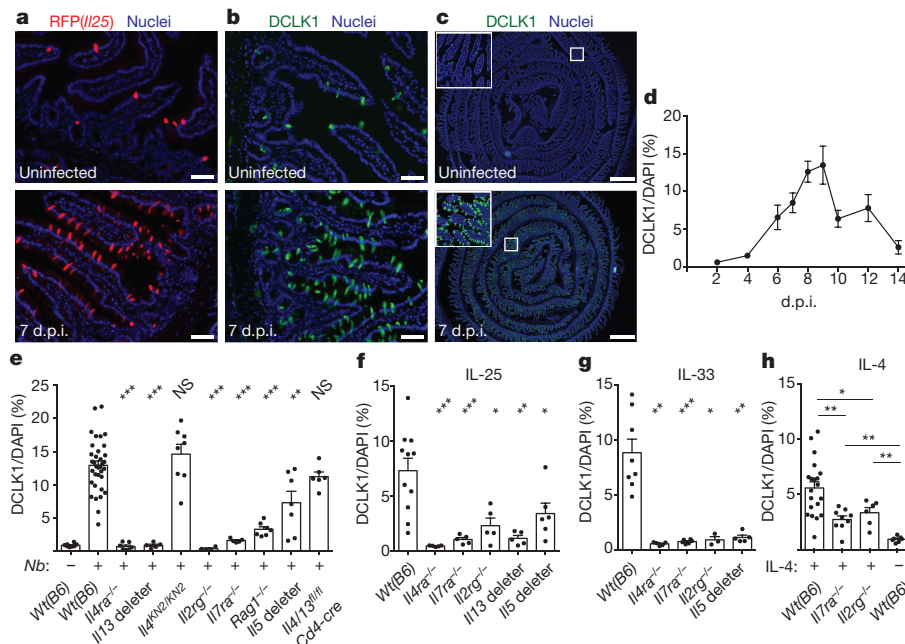


Figure 2 | Worm infection induces IL-13-dependent tuft cell hyperplasia. **a–c**, Jejunum from *Il25*^{F25/F25} mice stained for RFP (**a**) or DCLK1 (**b**, **c**). **d–h**, Immunohistochemical quantification of tuft cells (DCLK1⁺) in duodenum/jejunum of mice infected with *N. brasiliensis* for indicated days (**d**) or 7 days (**e**) or injected with indicated protein (**f–h**). Scale bars: 50 μ m (**a**, **b**), 1 mm (**c**). All data are biological replicates.

Data are representative of at least three (**a–c**) experiments or pooled (**d–h**) from multiple experiments. In **a–c**, $n > 10$; in **d**, day 2: $n = 2$; days 4, 12, 14: $n = 4$; day 9: $n = 5$; days 6, 8, 10: $n = 6$; day 7: $n = 8$; in **e–h**, n is as shown. Nb, *N. brasiliensis*. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant (Mann–Whitney test). Error bars represent mean \pm standard error of the mean (s.e.m.).

Extended Data Fig. 2). We also found RFP in epithelial cells of the trachea and gall bladder, but not in haematopoietic cells (Extended Data Figs 2 and 3a).

The small intestinal epithelium consists of a single cell layer continuously repopulated from stem cells in underlying crypts; cells progress up the villi and are sloughed into the lumen with a turnover of 3–5 days. Nascent progenitors proliferate in the transit amplifying region before fate commitment to become absorptive enterocytes or, less frequently, one of four secretory cell types: Paneth, enteroendocrine, goblet, or tuft^{12,13}. We tested whether Flare25 marks one or more secretory lineages. Immunohistochemistry showed no colocalization of RFP with the enteroendocrine marker chromogranin A (CHGA), the Paneth-cell markers lysozyme (LYZ)1 and LYZ2, or the goblet-cell marker mucin 2 (MUC2) (Fig. 1c and Extended Data Fig. 4a, b). Unexpectedly, expression of RFP and the tuft-cell markers doublecortin-like kinase 1 (DCLK1) and epithelial prostaglandin-endoperoxide synthase 1 (PTGS1) completely overlapped (Fig. 1d and Extended Data Fig. 4a, b). Transcriptional analysis comparing sorted RFP⁺EPCAM⁺ with RFP[−]EPCAM⁺ intestinal epithelium demonstrated *Il25* expression almost exclusively in RFP⁺ cells (Fig. 1e), and confirmed co-staining results (Fig. 1f and Extended Data Fig. 3b). The tuft-cell markers *Dclk1*, *Ptgs1*, *Gnat3*, *Chat*, *Gfi1b* and *Trpm5* (ref. 14, 15) were each enriched at least 750-fold in RFP⁺ cells, while *Chga*, *Muc2*, *Lyz1* and *Lyz2* showed no enrichment (Fig. 1f). Finally, >99% of sorted RFP⁺EPCAM⁺ and <1% of RFP[−]EPCAM⁺ cells were DCLK1⁺ by flow cytometry (Fig. 1g). Given these results, and our identification of RFP⁺ cells only in epithelia where tuft cells have been noted (Extended Data Figs 2 and 4c; data not shown)¹⁴, we conclude that tuft cells constitutively express *Il25* and that all *Il25*⁺ cells are tuft cells, at least as assessed using this reporter. By contrast, tuft cells are not major sources of TSLP or IL-33 in the small intestine (Extended Data Fig. 5).

Although intestinal tuft cells (also called brush cells) were discovered more than 50 years ago, their function remains largely unknown. Given the link between IL-25 and type 2 immunity, we investigated the role of tuft cells during infection of mice with the roundworm

Nippostrongylus brasiliensis, which induces a strong type 2 immune response that clears intestinal worms 7–10 days post-infection (d.p.i.). Although tuft cells account for <1% of intestinal epithelium in uninfected mice, we found dramatic (>15-fold) tuft cell hyperplasia in the small intestine 7 d.p.i. (Fig. 2a–c). The extent of hyperplasia was uniform throughout the small intestine (Extended Data Fig. 6), but we focused further experiments on the duodenum and jejunum, where *N. brasiliensis* resides. We observed no hyperplasia in the stomach or colon (Extended Data Fig. 6), through which the worms briefly transit. Hyperplasia in the small intestine peaked 8–9 d.p.i. and returned to near homeostatic levels by 14 d.p.i. (Fig. 2d). As in uninfected mice, RFP⁺ cells were CHGA[−], MUC2[−] and LYZ1/2[−], and DCLK1⁺ and PTGS1⁺ 7 d.p.i. (Extended Data Fig. 7). Given the complete overlap of RFP and DCLK1, we used these markers interchangeably in further experiments.

Since helminth-induced goblet cell hyperplasia is mediated by IL-13 (ref. 16), we asked whether IL-13 also mediates tuft cell hyperplasia. Indeed, tuft cell hyperplasia was absent in infected interleukin 4 receptor α (*Il4ra*)^{−/−} mice, in which both IL-4 and IL-13 signalling is disrupted, and in *Il13* deleter mice (*Il13*^{cre/cre}; *Gt(ROSA)26*^{STOP-flox::DTA/+}) (Fig. 2e). Because tuft cell hyperplasia was normal in IL-4-deficient mice (*Il4*^{KN2/KN2}) (Fig. 2e), and because the predominant role of IL-13 in IL-25-mediated pathologies is well established^{17,18}, we conclude that IL-13 is the primary signal driving tuft cell hyperplasia *in vivo*. We also found *Il4/13*-dependent tuft cell hyperplasia after infection with *Heligmosomoides polygyrus*, another intestinal parasite (Extended Data Fig. 7c).

Lamina propria ILC2s are the principal intestinal source of IL-13 in the first week of *N. brasiliensis* infection^{19,20}. Accordingly, tuft cell hyperplasia was absent or reduced 7 d.p.i. in mice lacking nearly all lymphoid cells (*Il7ra*^{−/−}, *Il2rg*^{−/−}) or IL-5⁺ ILC2 cells (*Il5* deleter: *Il5*^{cre/cre}; *Gt(ROSA)26*^{STOP-flox::DTA/STOP-flox::DTA}) (Fig. 2e). Hyperplasia was greatly reduced in *Rag1*^{−/−} mice, but nearly normal in *Il4/13*^{fl/fl}; *Cd4-cre* mice (Fig. 2e), consistent with the model that T cells produce little IL-13 at early time points but can boost IL-13 levels by supporting ILC2 activation²¹.

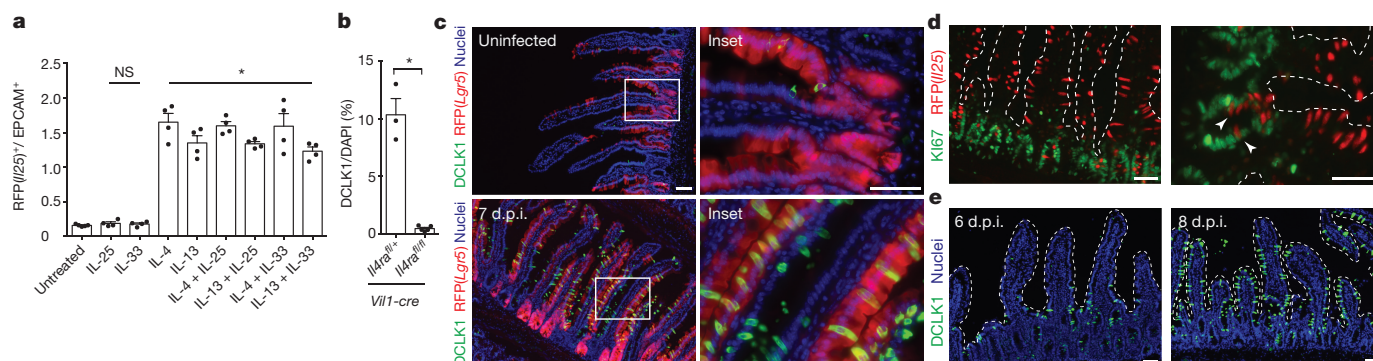


Figure 3 | IL-13 signalling in epithelial progenitors gives rise to tuft cell hyperplasia. **a**, Flow cytometric quantification of RFP⁺ tuft cells in intestinal organoids grown from *Il25*^{F25/F25} mice and treated with indicated recombinant proteins (20 ng ml⁻¹). **b**, Immunohistochemical quantification of tuft cells (DCLK1⁺) in duodenum/jejunum of mice infected 7 days with *N. brasiliensis*. **c**, Jejunum of *Lgr5*^{cre-Ert2/+}; *Gt(ROSA)26*^{STOP-flox::RFP/+} mice treated 5 days with tamoxifen and stained for DCLK1 (green) and DAPI (blue). *N. brasiliensis* infection as indicated. **d**, **e**, Jejunum of *Il25*^{F25/F25} (**d**) or *Wt(B6)* (**e**) mice infected for

7 days (**d**) or indicated number of days (**e**) with *N. brasiliensis* and stained for RFP (red) and Ki67 (green) (**d**) or DCLK1 (green) and DAPI (blue) (**e**). **d**, **e**, Dotted lines outline villi. Scale bars, 50 μ m. Arrowheads indicate Ki67 and RFP overlap. Data in **a** are technical replicates, all other data are biological replicates. Data are representative of two (**c–e**) or three (**a**) experiments or pooled (**b**) from multiple experiments. In **a**, **b**, **n** is as shown; in **c**, uninfected: *n* = 4; 7 d.p.i.: *n* = 5; in **d**, *n* = 2; in **e**, *n* = 4. *Nb*, *N. brasiliensis*. **P* < 0.05; NS, not significant (Mann–Whitney test). Error bars represent mean \pm s.e.m.

To test if type 2 cytokines are sufficient to induce tuft cell hyperplasia, we injected mice with stabilized IL-4–anti-IL-4 complexes that mimic IL-13 signalling, or with IL-25 or IL-33. All three treatments expanded tuft cells, but the effects of IL-25 and IL-33 were IL4RA-dependent and severely reduced in the absence of lymphoid cells (*Il7ra*^{-/-}, *Il2rg*^{-/-}), *Il5*⁺ cells, or *Il13*⁺ cells (Fig. 2f, g), suggesting that IL-25 and IL-33 trigger hyperplasia indirectly by inducing IL-13 production in ILC2s. By contrast, tuft cell hyperplasia was reduced only partially in *Il7ra*^{-/-} and *Il2rg*^{-/-} mice injected with IL-4 (Fig. 2h), raising the possibility that exogenous IL-4 or endogenous IL-13 induces tuft cell hyperplasia by directly targeting the epithelium.

Consistent with this model, recombinant IL-4 and IL-13 induced tuft cell hyperplasia in intestinal organoids, which contain only epithelial cells; by contrast, IL-25 and IL-33 neither induced hyperplasia nor enhanced induction by IL-4/IL-13 in this system (Fig. 3a and Extended Data Fig. 3c, d). As expected, tuft cell hyperplasia was absent in *Il4ra*^{fl/fl}; *Vil1-cre* mice, confirming that tuft cell hyperplasia requires IL4RA signalling in the intestinal epithelium *in vivo* (Fig. 3b). We conclude that ILC2-derived IL-13 signals through IL4RA in the intestinal epithelium to induce tuft cell hyperplasia.

We examined several possible mechanisms of IL-13-induced tuft cell hyperplasia. First, CHGA⁺ cell numbers did not change during infection (Extended Data Fig. 8b, c), confirming a selective increase in tuft cells rather than a global expansion of intestinal epithelium. Through lineage tracing, we confirmed that tuft cells, as under resting conditions, continue to arise from *Lgr5*⁺ stem cells during *N. brasiliensis*-induced hyperplasia (Fig. 3c). Since *Il4ra* is expressed throughout the intestinal epithelium^{22,23}, tuft cell expansion could occur either before or after lineage commitment; however, we found expression of the proliferation marker Ki67 in only a few nascent tuft cells (Fig. 3d), suggesting that tuft cell hyperplasia is induced in the stem or transit amplifying compartments. Consistent with this model, kinetic studies revealed a wave of hyperplastic tuft cells appearing near the crypts 6 d.p.i. and moving up the villi by 8 d.p.i. (Fig. 3e). Taken together, these results suggest that IL-13 signalling in uncommitted intestinal epithelium shifts cell fate decisions towards the tuft (and goblet) cell lineage, perhaps by altering the balance of Notch signalling¹². Indeed, the Notch signalling inhibitor *N*-[*N*-(3,5-Difluorophenacetyl)-L-alanyl]-S-phenylglycine *t*-butyl ester (DAPT) also induced tuft cell hyperplasia in organoids (Extended Data Fig. 3e–f).

Since ILC2s secrete IL-13 in response to IL-25, and tuft cells are the source of intestinal IL-25, we hypothesized a feed-forward circuit between tuft cells, ILC2s, and epithelial progenitors (Extended Data Fig. 8d). Tuft cells constitutively express *Il25* (Fig. 1a, e) and some

lamina propria ILC2s constitutively express *Il13* (ref. 24), so we first examined the interaction of tuft cells and ILC2s in uninfected mice. Using *Il13*^{Smart/Smart} reporter mice to mark IL-13-secreting cells, we found that ~20% of lamina propria ILC2s make IL-13 in the absence of infection, and this is dependent on IL-25 (Fig. 4a and Extended Data Fig. 9a). The frequency of ILC2s is also decreased in *Il25*^{-/-} mice (Fig. 4b), suggesting that tuft-cell-derived IL-25 promotes ILC2 maintenance in the small intestine. ILC2 activation remains IL-25-dependent at 4 d.p.i., the latest time point before worm clearance at which we could recover viable cells from infected intestines (Fig. 4a and Extended Data Fig. 9a). TH2 cells are not a major source of IL-13 at rest or 4 d.p.i. (Extended Data Fig. 9b, c).

Our model of an ILC2–epithelial circuit predicts that loss of homeostatic IL-25 or IL-13 would reduce the frequency of tuft cells. Indeed, the already small number of tuft cells was further reduced in uninfected *Il25*^{-/-} and *Il4ra*^{-/-} mice (Fig. 4c), with the remaining tuft cells in these mice probably representing stochastic production independent of type 2 immunity. The frequency of CHGA⁺ cells was unchanged in *Il25*^{-/-} mice (Extended Data Fig. 8b, c). We generated *Il25*^{F25/F25}; *Vil1-cre* mice to delete *Il25* selectively from the epithelium (Extended Data Fig. 1b, c). The basal frequency of tuft cells again decreased (Fig. 4d), confirming that tuft cells are the relevant source of IL-25 upstream of homeostatic production of IL-13 by intestinal ILC2s.

The dependence of tuft cell frequency on autocrine IL-25 was even more striking during *N. brasiliensis* infection. Tuft cell hyperplasia was absent in *Il25*^{-/-} and *Il25*^{F25/F25}; *Vil1-cre* mice 7 d.p.i., but unaffected by the absence of TSLP or IL-33 signalling (Fig. 4e, f). Taken together, our data support a model in which IL-25 from tuft cells induces ILC2s to produce IL-13, which in turn regulates the frequency of tuft cells in the intestinal epithelium. The capacity of IL-13 to alter the cellular composition of the epithelium raised the possibility that tuft cells might also regulate other secretory lineages. While neither CHGA⁺ nor Paneth cell hyperplasia occurred during worm infection (Extended Data Fig. 8a–c), goblet cell hyperplasia and hypertrophy were absent in infected mice lacking epithelial IL-25 (Fig. 4g–i). Moreover, as in *Il25*^{-/-} mice, *Il25*^{F25/F25}; *Vil1-cre* mice failed to clear worms by 10 d.p.i., thereby identifying tuft cells as key regulators of the type 2 immune response (Fig. 4j).

Our findings uncover an unexpected role for tuft cells in intestinal immune defence, suggest a link between type 2 immune signalling and epithelial cell fate decisions, and describe an ILC2–epithelial response circuit that regulates the cellular composition of the intestinal epithelium. This circuit could integrate homeostatic signals,

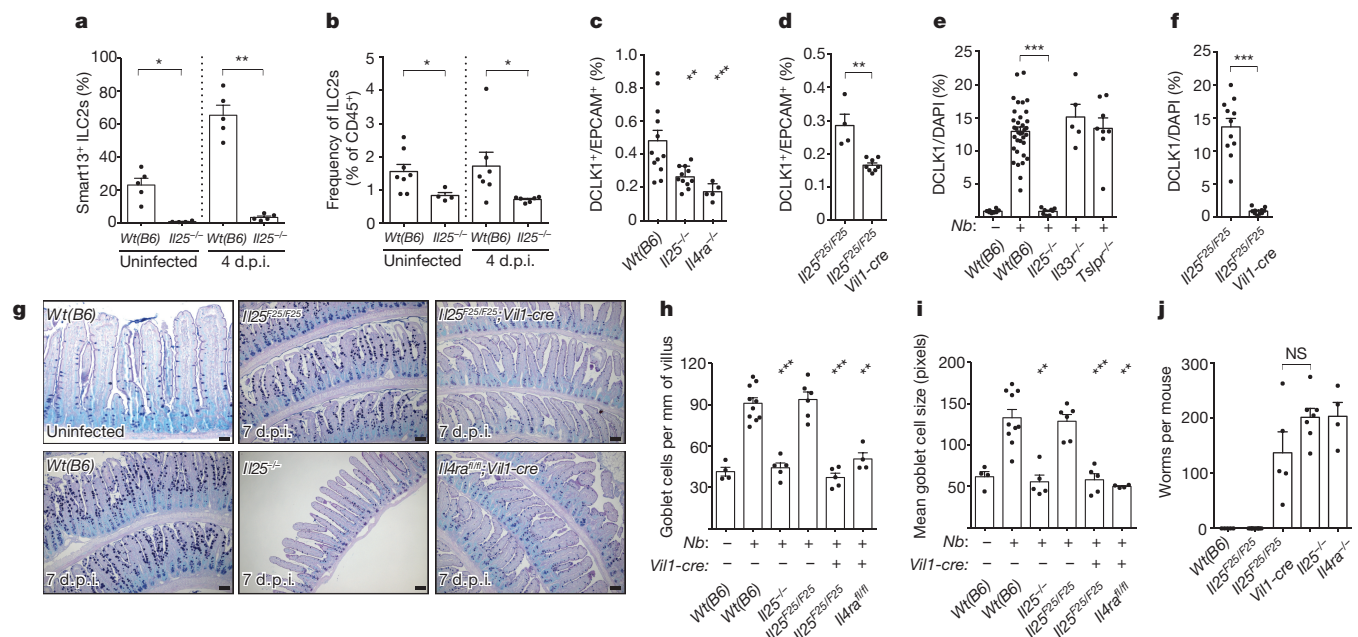


Figure 4 | Tuft cells regulate intestinal physiology through an ILC2–epithelium response circuit. **a, b**, Flow cytometric analysis of lamina propria ILC2s (Lin[−], CD45⁺, GATA3⁺) from *Il13^{Smart/+}* and *Il25^{−/−}*; *Il13^{Smart/+}* mice. **c, d**, Flow cytometric quantification of tuft cells (DCLK1⁺) in jejunum of uninfected mice. **e, f**, Immunohistochemical quantification of tuft cells (DCLK1⁺) in jejunum/duodenum of mice infected 7 days with *N. brasiliensis*. In **e**, wild-type controls are the same as in Fig. 2c. **g**, Jejunum of mice treated as indicated and stained for goblet

such as ILC2 activation during feeding²⁴, with additional signals to tune the barrier's absorptive–secretory balance. Indeed, despite the feed-forward nature of the circuit, tuft and goblet cell hyperplasia are restrained during homeostasis, suggesting that worm infection provides another activating signal or removes an inhibitory signal. Given their positioning in the intestinal epithelium, tuft cells appear poised to monitor luminal homeostasis and transduce activating signals to immune cells in the lamina propria. Interestingly, tuft cells encode the complete bitter and umami taste transduction pathways and release acetylcholine when activated¹⁴. In the airways and urethra, this pathway has been linked to smooth muscle contraction and is proposed to promote innate defence against invading bacteria^{25–28}.

In the absence of IL-25, other signals such as IL-33 become induced⁵ and can activate ILC2s to mediate expansion of tuft and goblet cells associated with worm clearance. Nonetheless, our findings delineate a key role for tuft cells in the physiological host response to helminths. Tuft cells appear to be the primary source of IL-25 in the lung as well (Extended Data Fig. 4c); thus, their involvement may extend to other conditions in which IL-25 has been implicated, such as airway disease^{11,29} and allergic diarrhoea³⁰. Indeed, mucosal tuft cell hyperplasia may be a generalizable hallmark of type 2 immune responses, and strategies to short-circuit tuft cell activation may have therapeutic potential for these widespread afflictions.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 June; accepted 21 October 2015.

Published online 14 December 2015.

1. Miller, H. R. & Nawa, Y. Immune regulation of intestinal goblet cell differentiation. Specific induction of non-specific protection against helminths? *Nouv. Rev. Fr. Hematol.* **21**, 31–45 (1979).
2. Castro, G. A., Badial-Aceves, F., Smith, J. W., Dudrick, S. J. & Weisbrodt, N. W. Altered small bowel propulsion associated with parasitism. *Gastroenterology* **71**, 620–625 (1976).

cells with periodic acid Schiff (PAS) Alcian blue. **h, i**, Goblet cell number (**h**) and size (**i**) calculated from imaging. **j**, Intestinal worm burden in mice infected 10 days with *N. brasiliensis*. Scale bars, 50 μ m. All data are biological replicates. Data are representative of two (**g**) experiments or pooled (**a–f, h–j**) from multiple experiments. In **g, n** is as shown in **h** and in **a–f, h–j, n** is as shown. *Nb, N. brasiliensis*. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; NS, not significant (Mann–Whitney test). Error bars represent mean \pm s.e.m.

3. Grecnis, R. K. Immunity to helminths: resistance, regulation, and susceptibility to gastrointestinal nematodes. *Annu. Rev. Immunol.* **33**, 201–225 (2015).
4. Price, A. E. *et al.* Systemically dispersed innate IL-13-expressing cells in type 2 immunity. *Proc. Natl Acad. Sci. USA* **107**, 11489–11494 (2010).
5. Neill, D. R. *et al.* Nuocytes represent a new innate effector leukocyte that mediates type-2 immunity. *Nature* **464**, 1367–1370 (2010).
6. Fallon, P. G. *et al.* Identification of an interleukin (IL)-25-dependent cell population that provides IL-4, IL-5, and IL-13 at the onset of helminth expulsion. *J. Exp. Med.* **203**, 1105–1116 (2006).
7. Kang, Z. *et al.* Epithelial cell-specific Act1 adaptor mediates interleukin-25-dependent helminth expulsion through expansion of Lin[−]c-Kit⁺ innate cell population. *Immunity* **36**, 821–833 (2012).
8. Zhao, A. *et al.* Critical role of IL-25 in nematode infection-induced alterations in intestinal function. *J. Immunol.* **185**, 6921–6929 (2010).
9. Van Dyken, S. J. *et al.* Chitin activates parallel immune modules that direct distinct inflammatory responses via innate lymphoid type 2 and $\gamma\delta$ T cells. *Immunity* **40**, 414–424 (2014).
10. Owyang, A. M. *et al.* Interleukin 25 regulates type 2 cytokine-dependent immunity and limits chronic inflammation in the gastrointestinal tract. *J. Exp. Med.* **203**, 843–849 (2006).
11. Angkasekwinai, P. *et al.* Interleukin 25 promotes the initiation of proallergic type 2 responses. *J. Exp. Med.* **204**, 1509–1517 (2007).
12. Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274–284 (2013).
13. Bjerknes, M. *et al.* Origin of the brush cell lineage in the mouse intestinal epithelium. *Dev. Biol.* **362**, 194–218 (2012).
14. Gerbe, F., Legraverend, C. & Jay, P. The intestinal epithelium tuft cells: specification and function. *Cell. Mol. Life Sci.* **69**, 2907–2917 (2012).
15. Bezençon, C. *et al.* Murine intestinal cells expressing Trpm5 are mostly brush cells and express markers of neuronal and inflammatory cells. *J. Comp. Neurol.* **509**, 514–525 (2008).
16. McKenzie, G. J., Bancroft, A., Grecnis, R. K. & McKenzie, A. N. J. A distinct role for interleukin-13 in Th2-cell-mediated immune responses. *Curr. Biol.* **8**, 339–342 (1998).
17. Fort, M. M. *et al.* IL-25 induces IL-4, IL-5, and IL-13 and Th2-associated pathologies *in vivo*. *Immunity* **15**, 985–995 (2001).
18. Hurst, S. D. *et al.* New IL-17 family members promote Th1 or Th2 responses in the lung: *in vivo* function of the novel cytokine IL-25. *J. Immunol.* **169**, 443–453 (2002).
19. Liang, H.-E. *et al.* Divergent expression patterns of IL-4 and IL-13 define unique functions in allergic immunity. *Nature Immunol.* **13**, 58–66 (2012).
20. Voehringer, D., Reese, T. A., Huang, X., Shinkai, K. & Locksley, R. M. Type 2 immunity is controlled by IL-4/IL-13 expression in hematopoietic non-eosinophil cells of the innate immune system. *J. Exp. Med.* **203**, 1435–1446 (2006).

21. Oliphant, C. J. *et al.* MHCII-mediated dialog between group 2 innate lymphoid cells and CD4⁺ T cells potentiates type 2 immunity and promotes parasitic helminth expulsion. *Immunity* **41**, 283–295 (2014).
22. Muñoz, J. *et al.* The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *EMBO J.* **31**, 3079–3091 (2012).
23. Reinecker, H. C. & Podolsky, D. K. Human intestinal epithelial cells express functional cytokine receptors sharing the common gamma c chain of the interleukin 2 receptor. *Proc. Natl Acad. Sci. USA* **92**, 8353–8357 (1995).
24. Nussbaum, J. C. *et al.* Type 2 innate lymphoid cells control eosinophil homeostasis. *Nature* **502**, 245–248 (2013).
25. Deckmann, K. *et al.* Bitter triggers acetylcholine release from polymodal urethral chemosensory cells and bladder reflexes. *Proc. Natl Acad. Sci. USA* **111**, 8287–8292 (2014).
26. Krasteva, G. *et al.* Cholinergic chemosensory cells in the trachea regulate breathing. *Proc. Natl Acad. Sci. USA* **108**, 9478–9483 (2011).
27. Saunders, C. J., Christensen, M., Finger, T. E. & Tizzano, M. Cholinergic neurotransmission links solitary chemosensory cells to nasal inflammation. *Proc. Natl Acad. Sci. USA* **111**, 6075–6080 (2014).
28. Lee, R. J. *et al.* Bitter and sweet taste receptors regulate human upper respiratory innate immunity. *J. Clin. Invest.* **124**, 1393–1405 (2014).
29. Ballantyne, S. J. *et al.* Blocking IL-25 prevents airway hyperresponsiveness in allergic asthma. *J. Allergy Clin. Immunol.* **120**, 1324–1331 (2007).
30. Han, H., Thelen, T. D., Comeau, M. R. & Ziegler, S. F. Thymic stromal lymphopoietin-mediated epicutaneous inflammation promotes acute diarrhea and anaphylaxis. *J. Clin. Invest.* **124**, 5442–5452 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Consengco, R. Noyes, and Z. Wang for technical expertise, Y. Nüsse for mice, members of the Locksley laboratory for helpful discussions, and R. Vance, M. Fontana, M. Anderson, and O. Klein for comments on the manuscript. This work was supported by the National Institutes of Health (AI026918, AI030663, HL107202), a Diabetes Endocrinology Research Center grant (DK063720), the Howard Hughes Medical Institute (HHMI), and the Sandler Asthma Basic Research Center at the University of California, San Francisco. J.v.M. is an HHMI Fellow of the Damon Runyon Cancer Research Foundation (DRG-2162-13).

Author Contributions J.v.M. conceived the study, designed and performed experiments, analysed data, and wrote the paper with R.M.L. M.J. performed experiments. H.-E.L. cloned the Flare25 reporter cassette, performed the Flare25 embryonic stem cell work, and assisted with additional experiments. R.M.L. directed the study and wrote the paper with J.v.M.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.M.L. (locksley@medicine.ucsf.edu).

METHODS

IL25 reporter mice. Flare25 mice were generated by homologous gene targeting in C57BL/6 embryonic stem cells. A 2.2 kb 3' homology arm beginning in the 3' UTR of *IL25* was amplified from C57BL/6 genomic DNA and cloned into pKO915-DT (Lexicon Genetics) using BamHI and HindIII. Next, a DNA strand encoding (from 5' to 3') a *loxP* site and the complete third exon of *IL25* was synthesized (Blue Heron). This synthetic strand and genomic DNA were used as PCR templates to generate a 2.1 kb 5' homology arm by overlap extension PCR. The 5' homology arm was cloned into the pKO915-DT vector containing the 3' homology arm using XhoI and EcoRI. Finally, a reporter cassette encoding (in order from 5' to 3'): a *loxP* site, encephalomyocarditis virus IRES, tandem RFP, bovine growth hormone poly(A), and a frt-flanked neomycin resistance cassette, was subcloned into the homology arm containing pKO915-DT vector using AscI. The final construct was linearized with NotI and transfected by electroporation into C57BL/6 embryonic stem cells. Cells were grown on irradiated feeders with the aminoglycoside G418 in the media, and neomycin-resistant clones were screened for 5' and 3' homologous recombination by PCR. Four positive clones were selected and further tested to confirm insertion of the 5' *loxP* site. Two clones were selected for injection into albino C57BL/6 blastocysts to generate chimaeras, and the male pups with highest ratios of black-to-white coat colour from a single clone were selected to breed with homozygous Gt(ROSA26)^{FLP1/FLP1} females (Jackson Laboratories catalogue no. 009086) to excise the neomycin resistance cassette. Deletion of neomycin was confirmed by PCR. Flare25 genotyping primers were as follows: KI_F: GTATTGGGTGCCAGAACAG; KI_R: GGGTCGCTACAGACGTTGTTTGTC (715 bp knock-in band; 374 bp floxed band); WT_F: ACTTTACCACAACCAGACG; WT_R: AGTTTCTCCCAAGTCCTCC (290 bp wild-type band).

Other mice. Mice were maintained in the University of California San Francisco (UCSF) specific pathogen-free animal facility in accordance with the guidelines established by the Institutional Animal Care and Use Committee and Laboratory Animal Resource Center. All experimental procedures were approved by the Laboratory Animal Resource Center at the UCSF. Mice aged 6–12 weeks were used for all experiments. Mice were age- and sex-matched in figures displaying a single representative experiment. Pooled results include both male and female mice. For some experiments, mice encoding a reporter allele that does not impact endogenous gene expression (B6.*Il13*^{Smart} or B6.*Arg1*^{YARG}) were used as wild-type controls. *Il7ra*^{−/−} (B6.129S7-*Il7r*^{tm1Imx/J}; 002295), *Rag1*^{−/−} (B6.129S7-*Rag1*^{tm1Mom/J}; 002216), *Cd4-cre* (B6.Cg-Tg(Cd4-cre); 022071) and wild-type (C57BL/6J; 000664) mice were purchased from Jackson Laboratories. *Il4ra*^{−/−} (BALB/c-*Il4ra*^{tm1Sz/J}; 003514) mice were purchased from Jackson Laboratories and backcrossed to C57BL/6J for at least eight generations. *Il2rg*^{−/−} (B10.B6-*Il2rg*^{tm1Wjl}; 4111-F) were purchased from Taconic as *Rag-2*^{−/−}; *Il2rg*^{−/−} and outcrossed to isolate the *Il2rg* allele. B6.*Il25*^{−/−}, B6.*Tslpr*^{−/−}, B6.*Il33*^{−/−}, B6.*Il13*^{Smart/Smart}, B6.*Il4*^{KN2/KN2}, B6.*Il5*^{cre/cre}; Gt(ROSA)26^{STOP-flox::DTA/STOP-flox::DTA} and B6.*Il13*^{cre/cre}; Gt(ROSA)26^{STOP-flox::DTA/+} mice were obtained or generated as described^{9,19,24}. BALB/c.*Il4/13*^{fl/fl} mice were generated as described²⁰ and backcrossed to C57BL/6J for at least eight generations. B6.*Il4ra*^{fl/fl} mice were provided by A. Chawla. B6.*Tg(Vil1-cre)* mice were provided by A. Ma. B6.*Lgr5*^{Egfp:cre-Ert2/+}; Gt(ROSA)26^{STOP-flox::RFP/+} mice were provided by O. Klein.

Mouse infection and treatment. Infectious third-stage *N. brasiliensis* larvae (L3) were raised and maintained as described¹⁹. Mice were infected subcutaneously with 500 *N. brasiliensis* L3 or by oral gavage with 200 *H. polygyrus* L3, and were killed at the indicated time points to collect tissues for staining or to count intestinal worm burden, as described¹⁹. Mice were given IL-4, IL-25, and IL-33 as follows: IL-4 complexes were generated by incubating 2 µg mouse IL-4 (R&D Systems) with 10 µg LEAF purified anti-mouse IL4 antibody (clone 11B11, Biolegend) for 30 min at room temperature, and then administered on day 0 and day 2. IL-25 and IL-33 were given in doses of 500 ng on days 0, 1, 2, and 3. All injections were given intraperitoneally in 200 µl, and all intestines were harvested for sectioning and staining on day 4. For lineage tracing, 2.5 mg of tamoxifen in 250 µl corn oil were given intraperitoneally 5 days before harvest.

Fixed tissue preparation and staining. For immunohistochemistry, tissues were fixed in 4% paraformaldehyde for 3 h at 4°C followed by PBS wash and overnight incubation in 30% (w/v) sucrose. For stomach, small intestine, caecum and large intestine, tissues were flushed with PBS before fixation. Unless otherwise noted, the proximal 10–12 cm of small intestine (duodenum and partial jejunum) were harvested. Tissues were embedded in Optimal Cutting Temperature Compound (Thermo-Tek) and stored at −80°C before sectioning (8–10 µm) on a Cryostat (Leica). To facilitate analysis of the entire sample, small and large intestines were coiled into a 'Swiss roll' before embedding.

Immunohistochemistry was performed in Tris/NaCl blocking buffer (0.1 M Tris-HCl, 0.15 M NaCl, 5 µg ml^{−1} blocking reagent (Perkin Elmer), pH 7.5) as follows: 1 h 5% goat serum, 1 h primary antibody, 40 min secondary antibody, 5 min

DAPI (Roche). For RFP co-labelling experiments, slides were stained for MUC2, LYZ1, CHRA, or DCLK1 as described above, excluding the DAPI step. RFP staining was then as follows: 1 h rabbit IgG, 20 min each of biotin and streptavidin block (Vector Labs), 1 h anti-RFP-biotin, 40 min streptavidin-Cy3 (Caltag), and 5 min DAPI. See Extended Data Table 1 for a list of antibodies used in this study.

For goblet cell staining, 8-cm sections of jejunum were fixed for 3 h in 10% buffered formalin (Fisher Scientific) at 4°C before coiling into a 'Swiss roll' and returning to formalin. After 24 h, tissues were moved to 70% ethanol for storage. Tissue processing, paraffin embedding, and sectioning were performed by the UCSF Mouse Pathology Core. Periodic acid Schiff (PAS) and Alcian blue staining were performed as follows: cleared with xylenes (Fisher Scientific), rehydrated, 30 min in Alcian blue (Thermo Scientific), 5 min in periodic acid (Thermo Scientific), 15 min in Schiff reagent (Thermo Scientific), dehydrated, and mounted. Brightfield and fluorescent images were acquired with an AxioCam HR camera on an AxioImagerM2 upright microscope (Zeiss).

Tuft and goblet cell quantification. In uninfected mice, a 2.5-cm section of small intestine was harvested beginning 10 cm distal to the stomach and processed into a single-cell epithelial suspension as described later. After analysis by flow cytometry, frequency of tuft cells was calculated as number of DCLK1⁺EPCAM⁺ cells/total number of EPCAM⁺ cells. Because viable epithelial cells cannot be harvested from *N. brasiliensis*-infected intestines from ~5 to ~12 d.p.i., immunohistochemistry was used to quantify tuft cell frequency in infected mice. The proximal 10 cm of small intestine were harvested and stained for DCLK1 as described earlier. A 4 × 4 grid of images was collected at ×200 magnification and the total area of DCLK1 and DAPI staining above background was calculated using ImageJ. Tuft cell frequency was calculated as DCLK1 staining area/DAPI staining area.

For goblet cell quantification, tissue sections were prepared and stained with PAS Alcian blue as described earlier. Goblet cells were manually counted and the total length of all analysed villi was measured using ImageJ. Goblet cell frequency was expressed as number of goblet cells/millimetre of villus. At least 15 villi were counted for each replicate. Mucus production was estimated by measuring the area of at least 15 goblet cells for each biological replicate.

Single-cell tissue preparation. For single-cell epithelial preparations, small intestines were flushed with PBS, opened, and rinsed with PBS to remove luminal contents. Two-and-a-half- to five-cm-long segments of jejunum were incubated with rocking for 20 min at 37°C in 5 ml PBS containing 2.5 mM EDTA (Sigma-Aldrich), 0.75 mM dithiothreitol (DTT; Sigma-Aldrich), and 10 µg ml^{−1} DNaseI (Sigma-Aldrich). Tissues were shaken vigorously for 30 s and released cells were incubated with rocking for 10 min at 37°C in 5 ml HBSS (Ca²⁺/Mg²⁺ free) containing 1.0 U ml^{−1} Dispase (Gibco) and 10 µg ml^{−1} DNaseI. Digested cells were passed through a 70 µm filter and washed once before staining for flow cytometry.

For lamina propria preparations, small intestine was harvested from 4–10 cm distal to the stomach (duodenum/jejunum), flushed with PBS, opened, and thoroughly cleaned with PBS. Intestines were incubated with gentle rocking for 15 min at 37°C in 10 ml HBSS (Ca²⁺/Mg²⁺ free) supplemented with 5% fetal calf serum (FCS), 10 mM HEPES (UCSF Cell Culture Facility), 10 mM DTT and 5 mM EDTA. Intestines were gently vortexed, supernatants discarded, and incubation repeated with fresh DTT/EDTA solution. Next, intestines were incubated with gentle rocking for 20 min at 37°C in 20 ml HBSS (Ca²⁺/Mg²⁺ replete) supplemented with 5% FCS and 10 mM HEPES. After incubation, intestines were gently vortexed, cut into small pieces and incubated with gentle rocking for 30 min at 37°C in 5 ml HBSS (Ca²⁺/Mg²⁺ replete) supplemented with 5% FCS, 10 mM HEPES, 30 µg ml^{−1} DNaseI, and 0.1 Wünsch ml^{−1} LiberaseTM (Roche). After digest, intestines were mechanically dissociated in GentleMACS C tubes (Miltenyi Biotec), passed through a 70 µm filter, and washed. The resulting cell pellet was resuspended in 4 ml 40% Percoll (Sigma-Aldrich), underlaid with 4 ml 90% Percoll and centrifuged at 2,200 r.p.m. for 20 min at 4°C. The 40/90 interphase of the Percoll gradient was harvested, washed, and stained for flow cytometry.

Flow cytometry. For surface staining, single-cell suspensions prepared as described earlier were incubated with anti-CD16 and CD32 monoclonal antibodies (UCSF Antibody Core Facility) for 10 min at 4°C. The cells were stained with antibodies to surface markers for 20 min at 4°C followed by DAPI for dead cell exclusion. See Extended Data Table 1 for a list of antibodies used in this study.

For intracellular DCLK1 staining, single-cell epithelial suspensions were prepared from 2.5 cm sections of small intestine harvested 8 cm distal to the stomach. Staining was as follows: 15 min at 4°C in Violet Live/Dead fixable stain (Life Technologies), 15 min at room temperature in 2% paraformaldehyde (Electron Microscopy Sciences), 20 min at room temperature in saponin-based permeabilization and wash (perm/wash) reagent (Life Technologies) supplemented with 10% goat serum, 30 min with rabbit anti-doublecortin-like kinase (Abcam; ab31704) in perm/wash, 20 min in F(ab')₂ goat anti-rabbit IgG-Alexa Fluor 488 (Life Technologies) and rat anti-EPCAM-PerCP-Cy5.5 (Biolegend; 16A8).

For intracellular GATA3 staining, single-cell lamina propria suspensions were prepared from the small intestines as described earlier and stained according to manufacturer's protocol for FoxP3/Transcription Factor Staining Buffer Set (eBiosciences).

Samples were analysed on an LSR II (BD Biosciences) with four lasers (403 nm, 488 nm, 535 nm, and 633 nm). Samples were FSC-A/SSC-A gated to exclude debris, FSC-W/FSC-A gated to select single cells, and gated to exclude DAPI⁺ dead cells. Data were analysed with FlowJo 10 (Treestar).

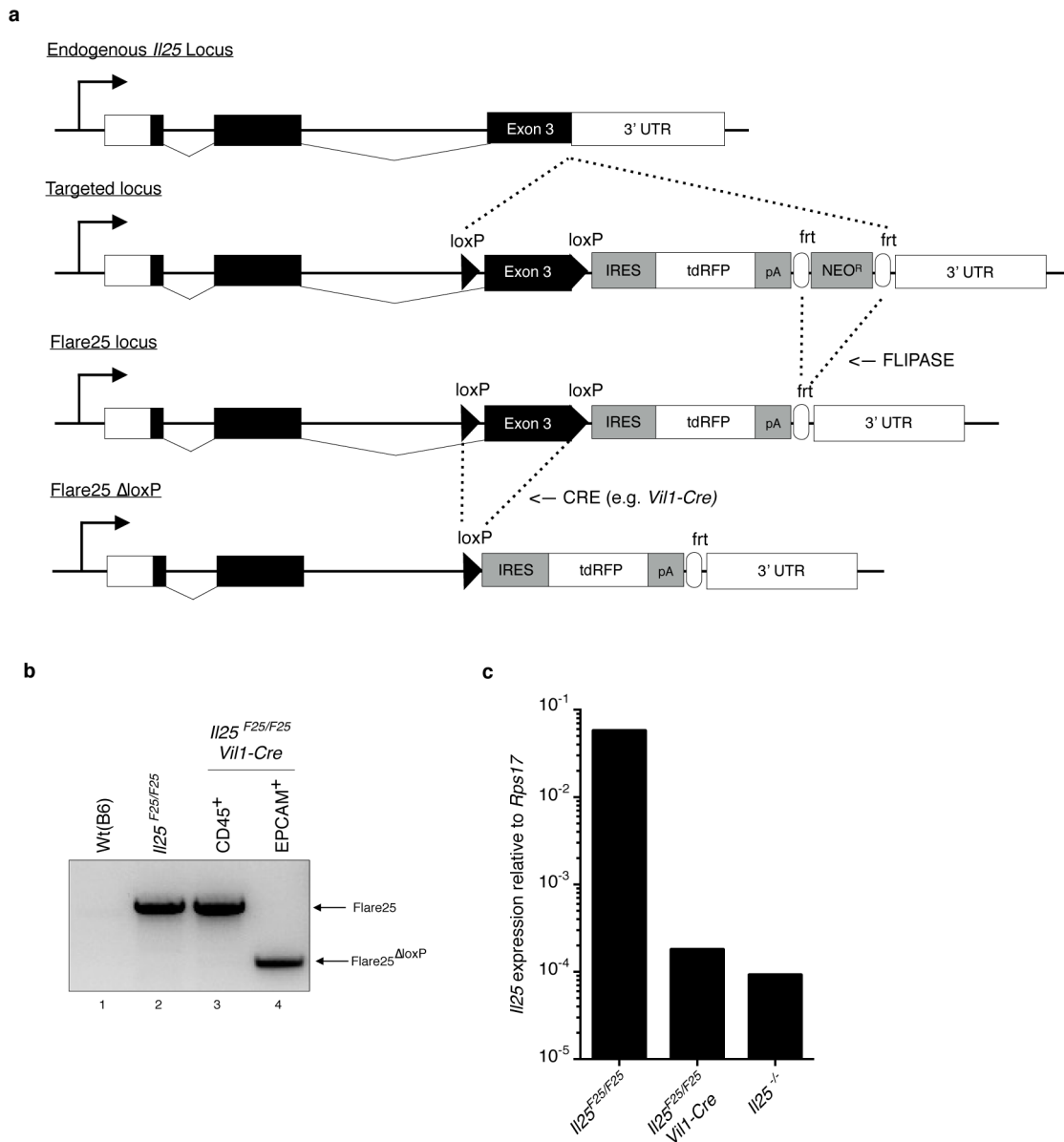
Quantitative RT-PCR. Single-cell epithelial suspensions were isolated and stained as described earlier and then sorted into RFP⁺EPCAM⁺ and RFP⁺EPCAM⁺ populations using a MoFlo XDP (Beckman Coulter). RNA was isolated using the Micro Plus RNeasy kit (Qiagen) and reverse transcribed using the SuperScript Vilo Master Mix (Life Technologies). The resulting cDNA was used as template for quantitative PCR with the Power SYBR Green reagent on a StepOnePlus cycler (Applied Biosystems). Transcripts were normalized to *Rps17* (40S ribosomal protein S17) expression. See Extended Data Table 1 for a list of primers used in this study.

Organoid culture. Small intestinal crypt-derived organoids were grown as described³¹, replacing recombinant R-spondin with supernatants from R-spondin expressing L-cells (provided by O. Klein). Crypts were harvested from *Il25*^{F25/F25}

mice and plated on day 0. On day 3 and day 5, media were replaced and organoids were treated with 20 ng ml⁻¹ of the indicated recombinant protein. On day 6 organoids were harvested into HBSS (Ca²⁺/Mg²⁺ replete) containing 200 U ml⁻¹ Collagenase I (Gibco) and 1.8 U ml⁻¹ Dispase (Gibco). Organoids were incubated for 1.5 h at 37 °C with shaking, washed, and then stained for flow cytometry as described earlier.

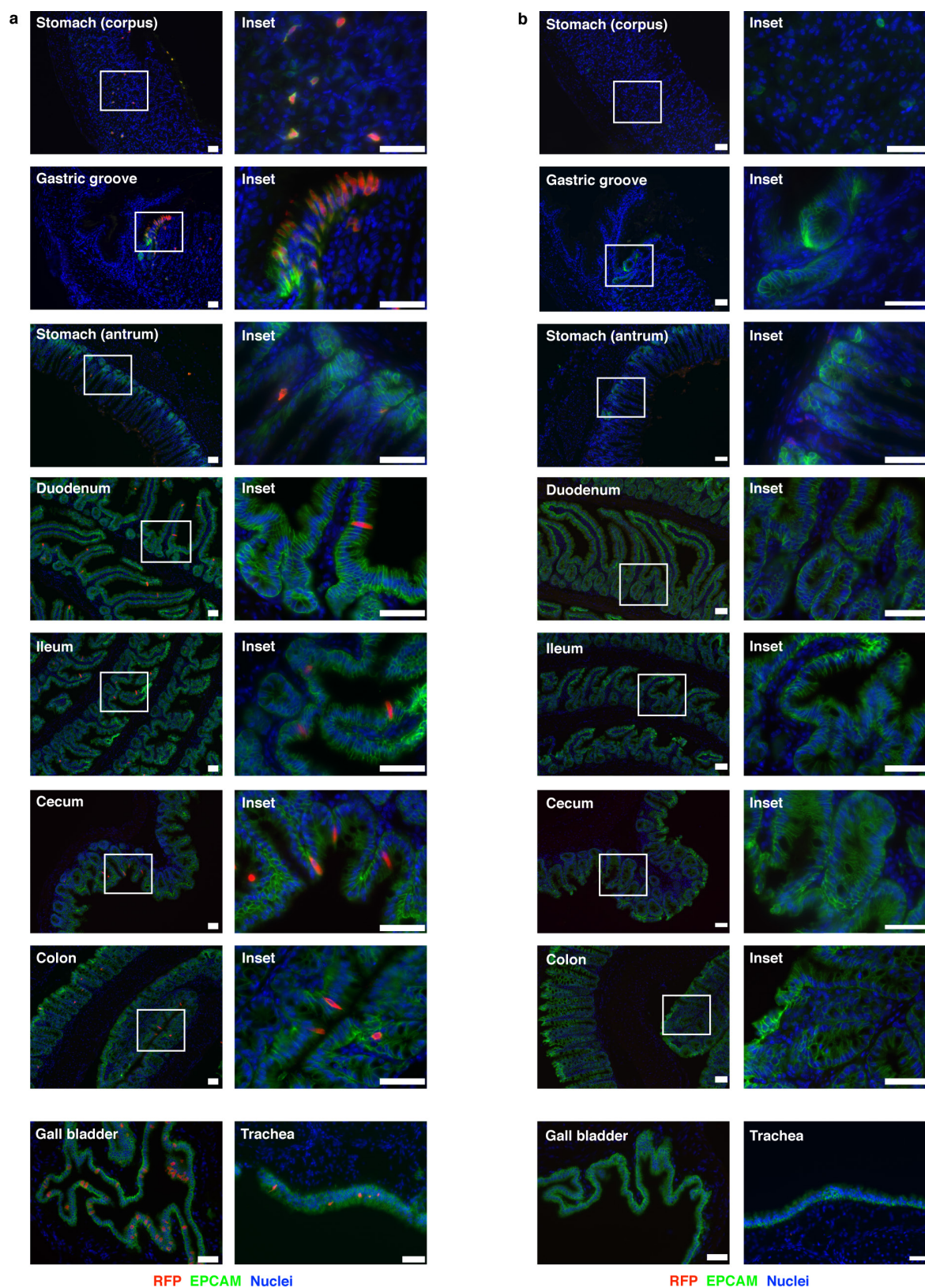
Statistical analysis. All experiments were performed using randomly assigned mice without investigator blinding. All data points and *n* values reflect biological replicates (that is, mice), except in Fig. 3a, where data points on the graph represent technical replicates. No data were excluded. Where noted in the figures, statistical significance was calculated without assumption of normal distribution using a Mann–Whitney test. Experimental groups included a minimum of three biological replicates, as required by the Mann–Whitney test. Intragroup variation was not assessed. All statistical analysis was performed using Prism 6 (GraphPad Software). Figures display means ± s.e.m. No statistical methods were used to predetermine sample size.

31. Sato, T. & Clevers, H. Primary mouse small intestinal epithelial cell cultures. *Methods Mol. Biol.* **945**, 319–328 (2013).



Extended Data Figure 1 | Flare25 mouse and *Vil1-cre*-mediated *Il25* deletion. **a**, Gene-targeting strategy for the flox and reporter of *Il25* (Flare25) mouse. **b**, PCR of genomic DNA isolated from the tail (lane 1, 2) or cells sorted from the small intestine (lane 3, 4) of indicated mice. **c**, Quantitative RT-PCR for *Il25* on cDNA from EPCAM⁺ cells sorted from the small intestine of indicated mice. **b**, **c**, Data are representative

of two experiments ($n = 2$). Frt, target site for FLIPASE recombinase; IRES, internal ribosomal entry site; *loxP*, target site for Cre recombinase; pA, bovine growth hormone poly(A) tail; tdRFP, tandem-dimer red fluorescent protein; UTR, untranslated region. For gel source data (**b**) see Supplementary Fig. 1.

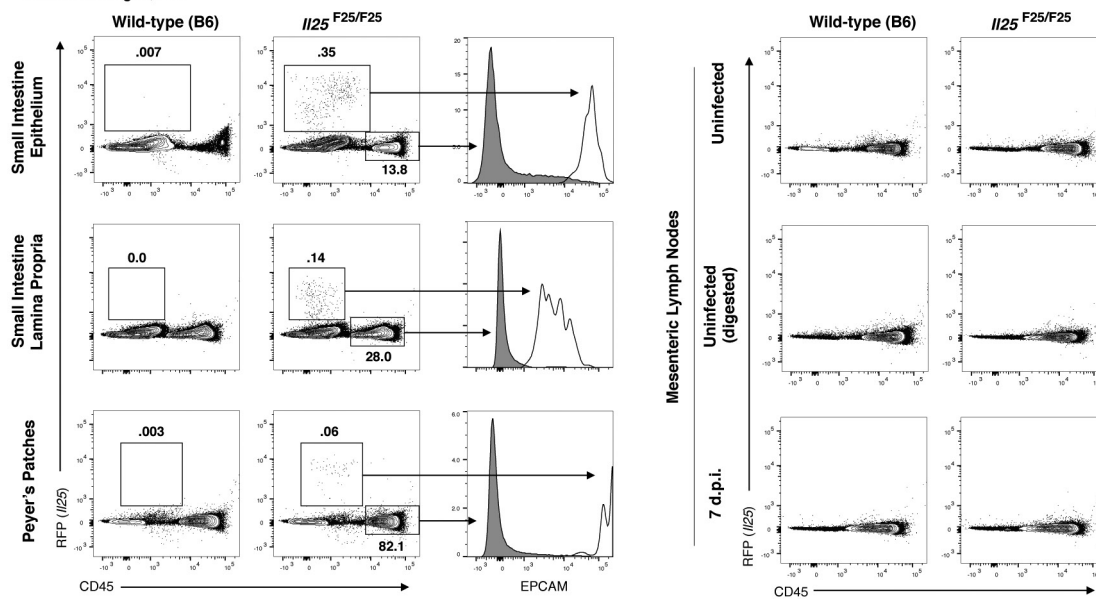


Extended Data Figure 2 | *IL25* expression in epithelial surfaces.
a, b, Indicated tissues of *IL25^{F25/F25}* (a) and wild-type control (b) mice stained by immunohistochemistry for RFP (red), EPCAM (green), and

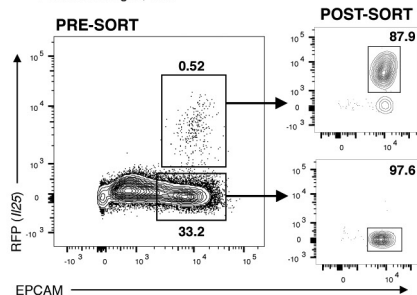
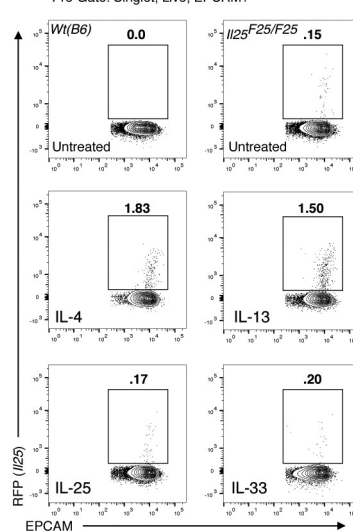
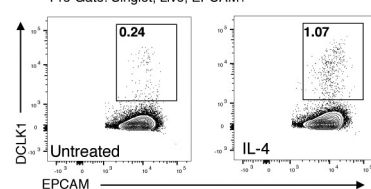
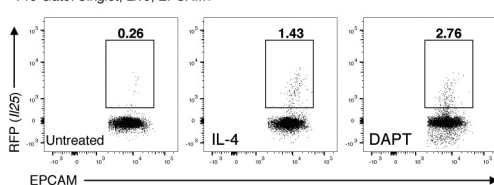
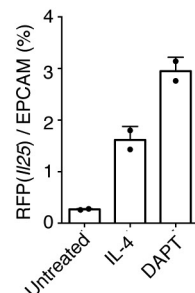
DAPI (blue). Some data from Fig. 1a are repeated here to allow complete comparison. Scale bars, 50 μm. Images are representative of at least three independent experiments. *n* = 3.

a

Pre-Gate: Singlet; Live

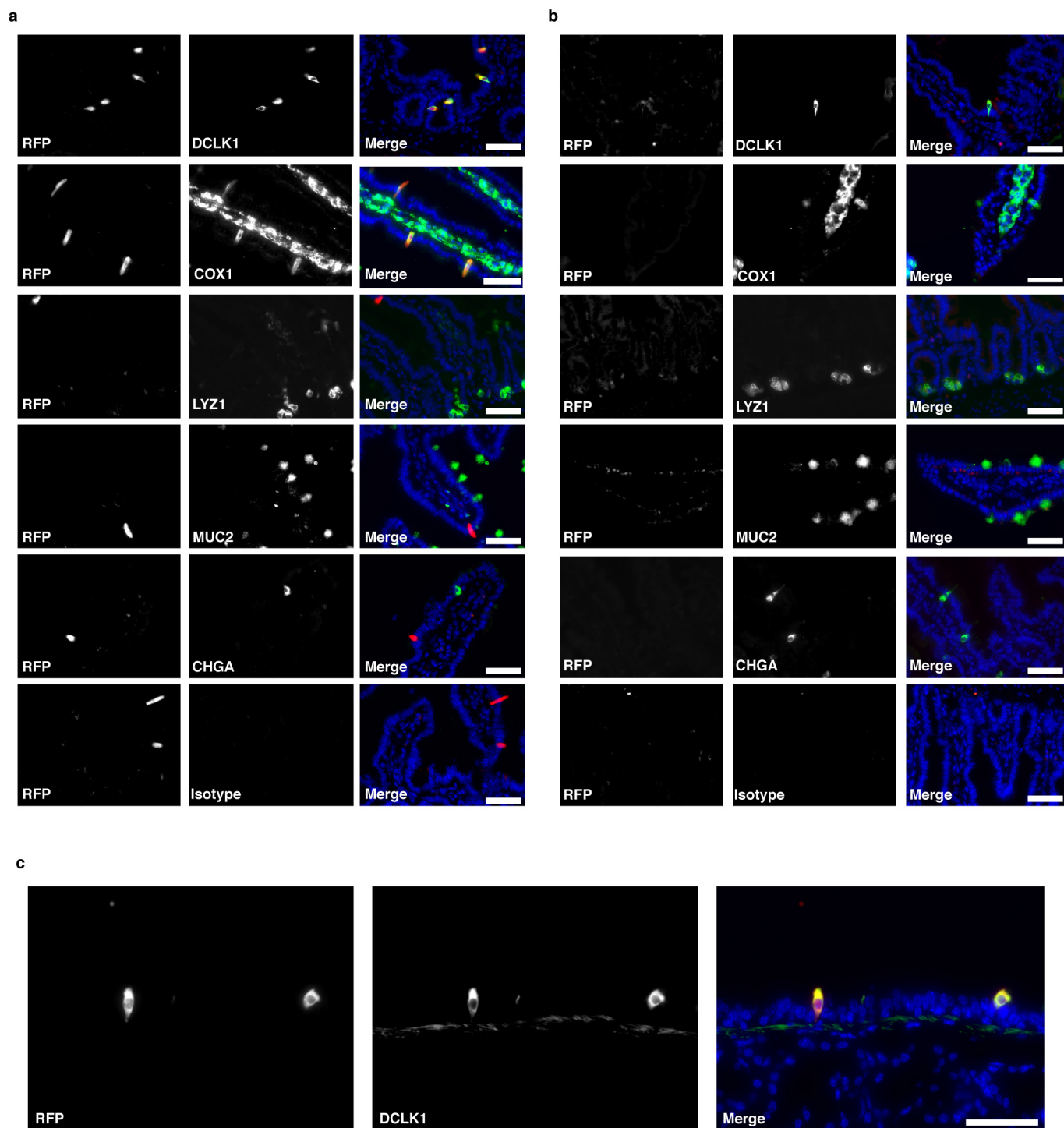
**b**

Pre-Gate: Singlet; Live

**c**Pre-Gate: Singlet; Live; EPCAM⁺**d**Pre-Gate: Singlet; Live; EPCAM⁺**e**Pre-Gate: Singlet; Live; EPCAM⁺**f**

Extended Data Figure 3 | Flow cytometry gating strategies and organoid culture. **a**, Flow cytometric analysis of indicated tissues in *Il25*^{F25/F25} and wild-type mice. **b**, Flow cytometric analysis of small intestine epithelial cells of *Il25*^{F25/F25} mice before and after fluorescence-activated cell sorting (FACS) into RFP⁺EPCAM⁺ and RFP⁺EPCAM⁺ pools for analysis by quantitative RT-PCR. **c–e**, Representative flow cytometric analysis of small-intestine-derived organoids from *Il25*^{F25/F25} (**c–e**) and wild-type (**c**) mice cultured with or without recombinant protein (20 ng ml⁻¹),

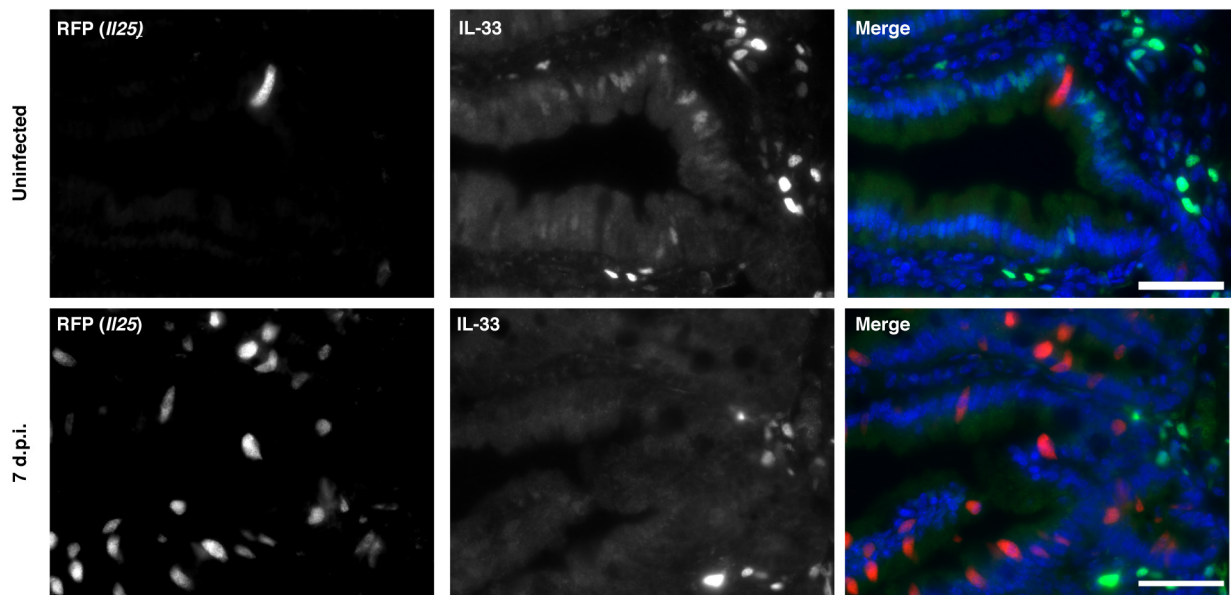
as indicated (**c–e**) or Notch signalling inhibitor DAPT (25 μM) (**e**). Single-cell suspensions of the organoids were stained for EPCAM (**c–e**) and DCLK1 (**d**), and gated to quantify tuft cell (RFP⁺EPCAM⁺ or DCLK1⁺EPCAM⁺) frequency. **f**, Quantification of two technical replicates from experiment shown in **e**. d.p.i., days post-*N. brasiliensis* infection. Data in **f** are technical replicates. Data are representative of three (**a**, **b**, **d**) or two (**c**, **e**, **f**) independent experiments. In **a–d**, *n* = 3; in **e**, **f**, *n* = 2. Error bars represent mean ± s.e.m.



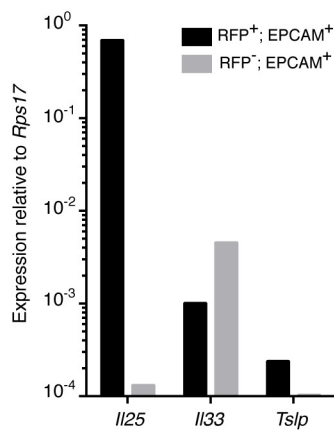
Extended Data Figure 4 | *Il25* is expressed constitutively in tuft cells.
a–c, Jejunum (a, b) or trachea (c) of *Il25^{F25/F25}* (a, c) and wild-type control (b) mice stained by immunohistochemistry for RFP (red), indicated

lineage markers (green), and DAPI (blue). Scale bars, 50 μ m. Images are representative of one (c) or two (a, b) independent experiments. $n = 2$.

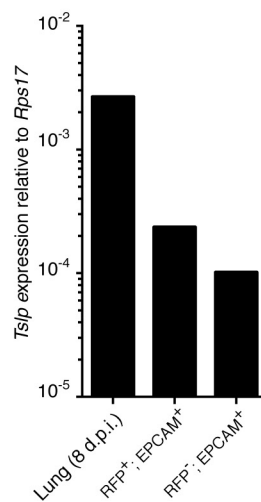
a



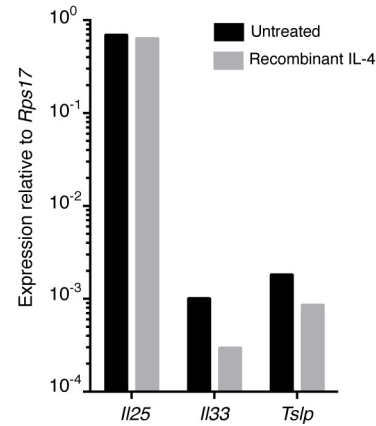
b



c

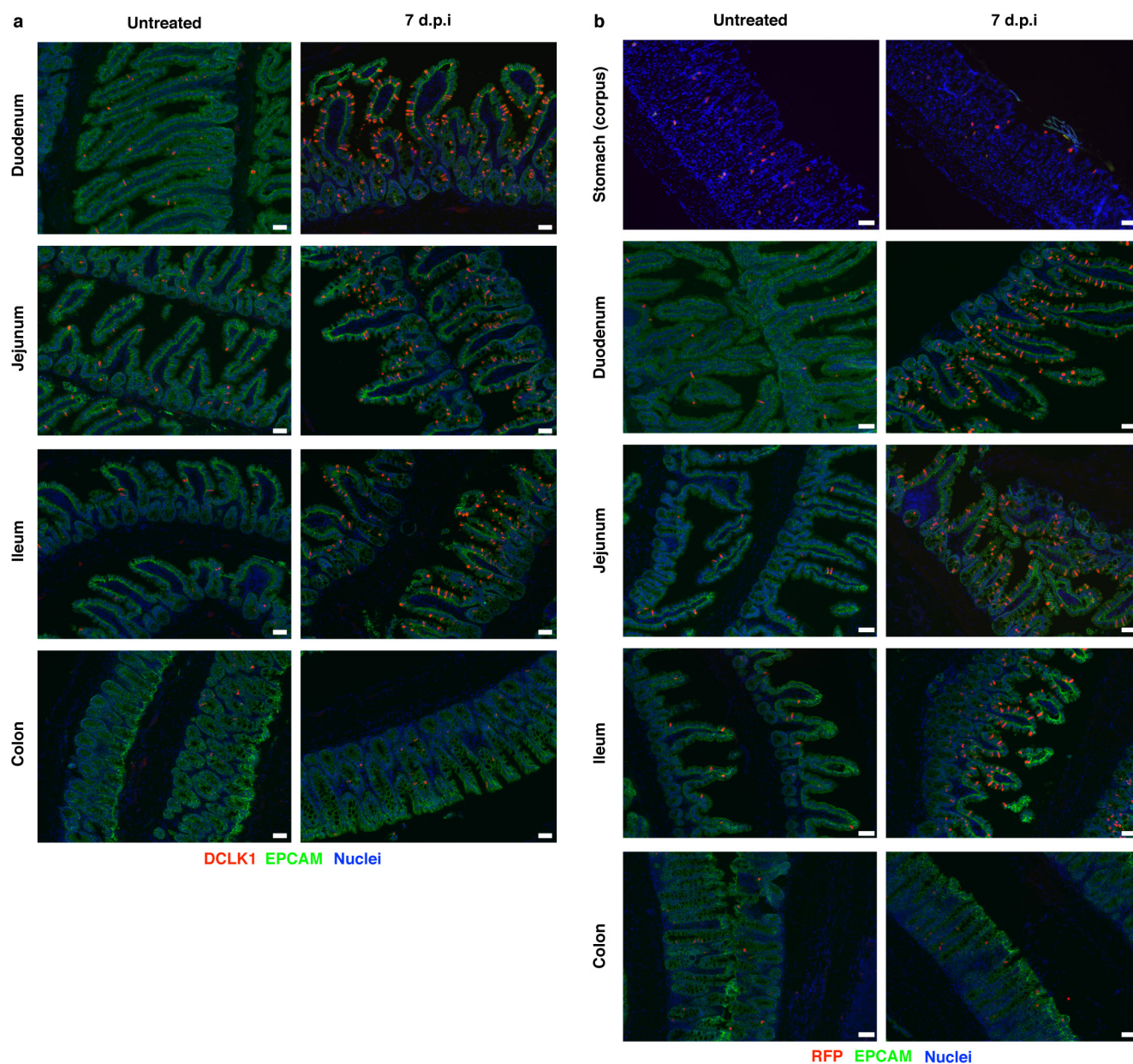


d



Extended Data Figure 5 | Tuft cells are not a major source of intestinal TSLP or IL-33. a, Jejunum of *Il25*^{F25/F25} mice stained for RFP (red), IL-33 (green), and DAPI (blue). **b–d,** Quantitative RT-PCR on indicated (**b, c**) or *RFP*⁺*EPCAM*⁺ (**d**) cells sorted from untreated (**b, c**) mice or mice treated as indicated (**d**). RNA isolated from whole lung 8 days post-*N. brasiliensis*

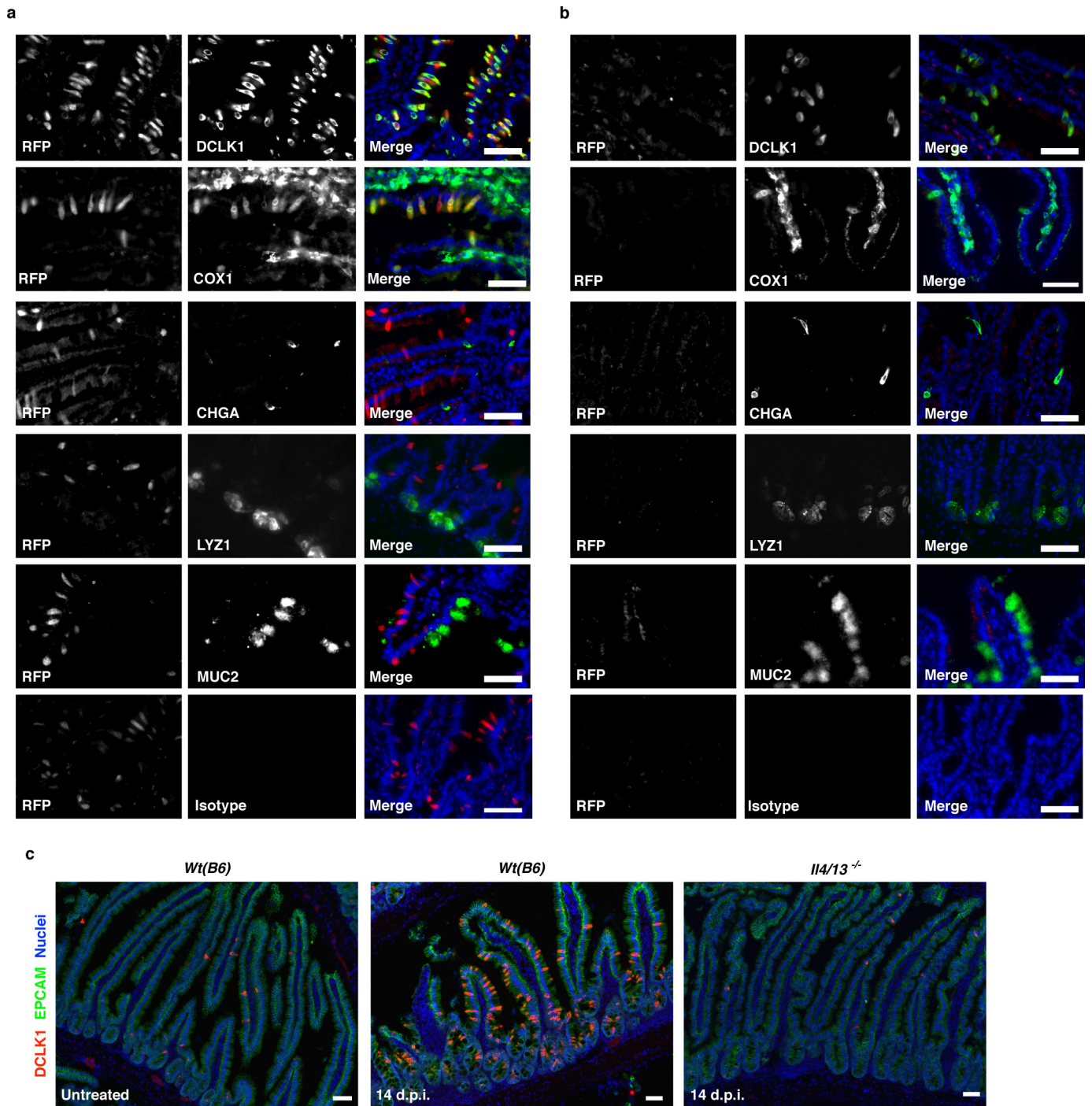
infection is used as a positive control for *Tslp* expression in **c**. Expression of *Tslp* in sorted *Tslp*-expressing cells of the lung would probably be higher. Scale bars, 50 μ m. Data are representative of two independent experiments. In **a**, $n = 3$; in **b–d**, $n = 2$.



Extended Data Figure 6 | *N. brasiliensis* induces tuft cell hyperplasia throughout the small intestine but not in stomach and colon.

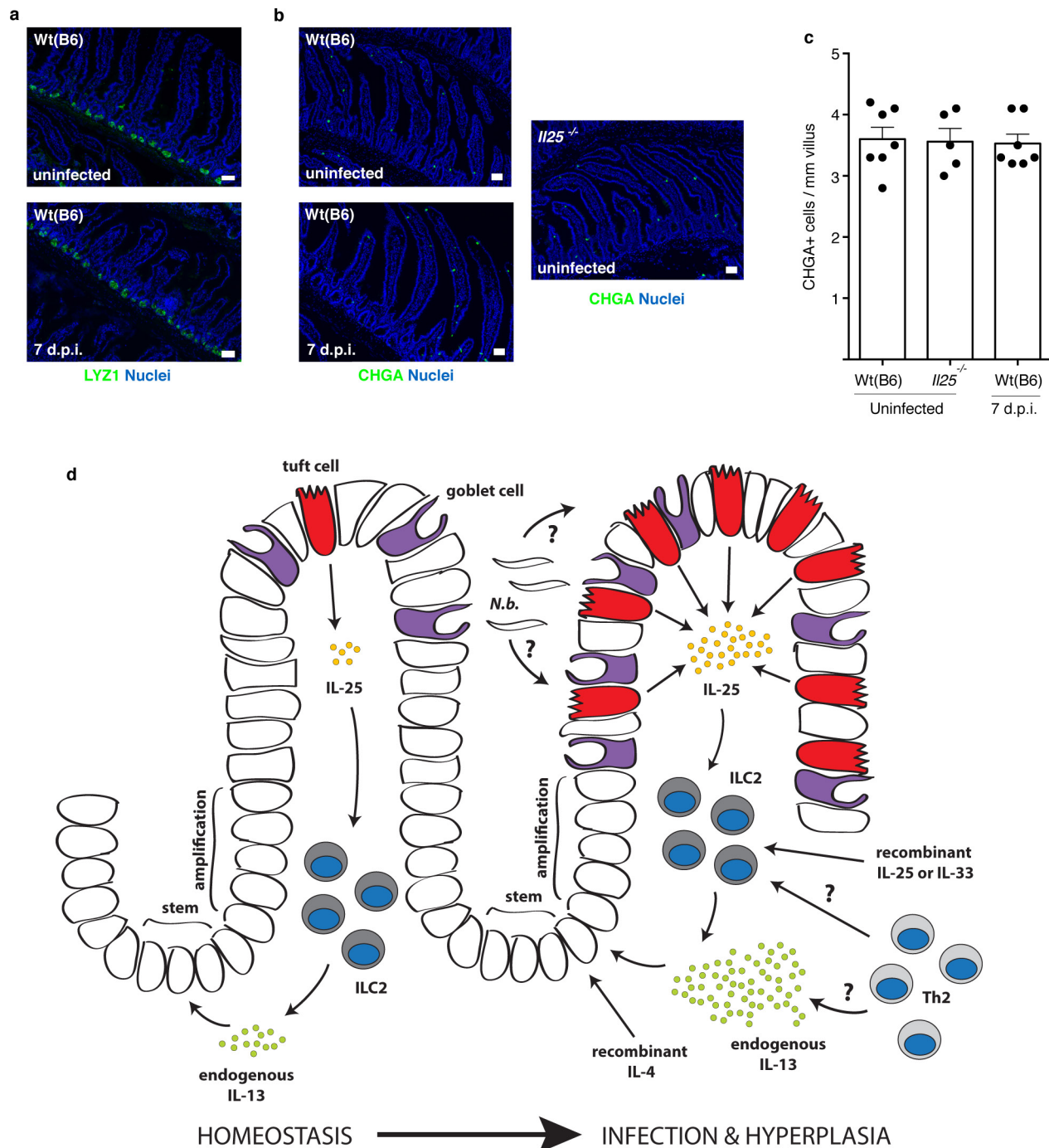
a, b, Indicated tissues of *Il25^{F25/F25}* (**a**) and wild-type control (**b**) mice treated as indicated and stained by immunohistochemistry for RFP (**a**) or

DCLK1 (**b**) (red), EPCAM (green), and DAPI (blue). d.p.i., days post-*N. brasiliensis* infection. Scale bars, 50 μ m. Data are representative of two (stomach and colon) or at least three (all others) independent experiments. In **a**, stomach and colon: $n = 2$; all others: $n > 5$.



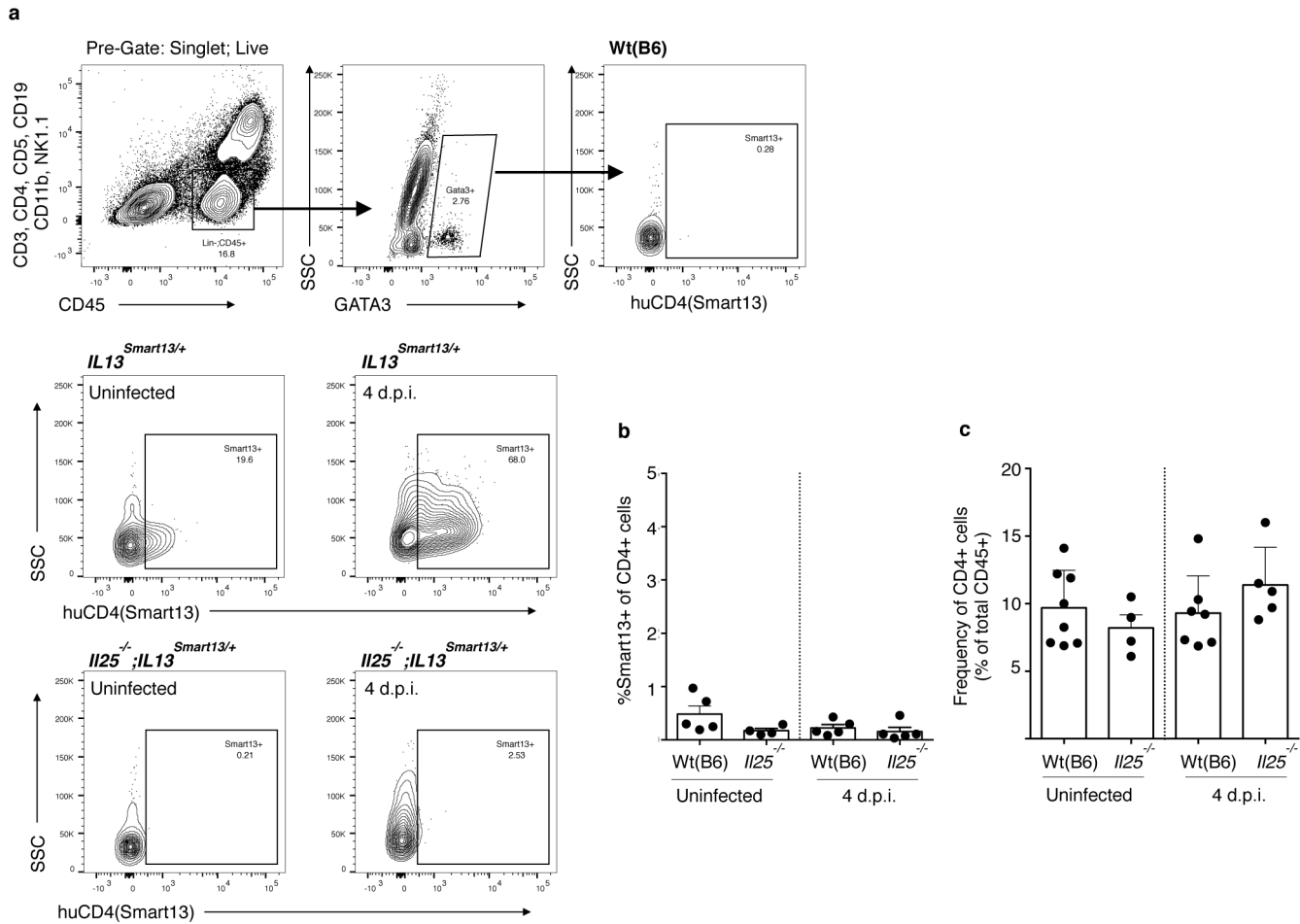
Extended Data Figure 7 | *Il25* is expressed only in tuft cells during worm infection and *H. polygyrus* infection also induces tuft cell hyperplasia. **a, b,** Jejunum of *Il25^{F25/F25}* (a) and wild-type control (b) mice infected for 7 days with *N. brasiliensis* stained by immunohistochemistry for RFP (red), indicated lineage markers (green), and DAPI (blue). **c,** Jejunum of indicated mice left untreated or infected 14 days with

H. polygyrus and stained by immunohistochemistry for DAPI (blue), EPCAM (green) and DCLK1 (red). Scale bars, 50 μm. d.p.i., days post-*H. polygyrus* infection. Images are representative of one (c) or two (a, b) independent experiments. In a, b, *n* = 2; in c, *n* = 1 (uninfected) or *n* = 2 (infected).



Extended Data Figure 8 | Absence of Paneth and CHGA⁺ cell hyperplasia after *N. brasiliensis* infection and model of ILC2-epithelial signalling circuit. **a, b,** Jejunum of indicated mice stained for DAPI (blue) and LYZ1/2 (**a**) or CHGA (**b**) (green). **c,** Quantification of CHGA⁺ cells from imaging in (**b**). **d,** During homeostasis, rare epithelial tuft cells of the small intestine constitutively express *Il25*, which maintains low levels of IL-13 production in lamina propria ILC2s. IL-13 in turn signals uncommitted epithelial progenitors to promote emergence of tuft and goblet cells. In the absence of infection, this feed-forward ILC2-epithelial circuit is restrained by as yet unknown mechanisms. After *N. brasiliensis* (*N.b.*) infection, a helminth-derived signal or a change in

host physiology activates the ILC2-epithelial circuit leading to tuft and goblet cell hyperplasia and enhanced IL-13 production by ILC2s. Adaptive Th2 cells probably also provide IL-13 and/or support ILC2 activation, especially when infection or inflammation lasts more than a week. Recombinant proteins are sufficient to induce tuft cell hyperplasia, either by inducing IL-13 production in lymphoid cells (IL-25 or IL-33) or by directly binding epithelial progenitors (IL-4). Scale bars, 50 μ m. d.p.i., days post-*N. brasiliensis* infection. Data in **c** are biological replicates. Data are representative of two (**a**) or three (**b**) independent experiments or pooled from multiple experiments (**c**). In **a**, $n = 2$; in **b**, n is as shown in **c**. Error bars represent mean \pm s.e.m.



Extended Data Figure 9 | IL-13 production by lamina propria ILC2 and CD4⁺ cells. **a, b**, Lamina propria cells from *Il25*^{-/-}; *Il13*^{Smart13/+}, *Il13*^{Smart13/+}, and wild-type control mice analysed by flow cytometry and gated on ILC2 (**a**, Lin⁻CD45⁺GATA3⁺) or CD45⁺CD4⁺ (**b**) cells. IL-13 secretion was quantified by measuring surface expression of human CD4, which is expressed from the *Il13*^{Smart} reporter mice. **c**, Frequency of

lamina propria CD4⁺ cells as a percentage of total CD45⁺ cells as assessed by flow cytometry. d.p.i., days post-*N. brasiliensis* infection. Data in **b, c** are biological replicates. Data are representative of at least three (**a**) independent experiments, or pooled from multiple experiments (**b, c**). In **a**, *n* = 5; in **b, c**, *n* is as shown. Error bars represent mean ± s.e.m.

Extended Data Table 1 | Antibodies and quantitative RT-PCR primers used in this study

a					b				
Immunohistochemistry Antibodies					Flow Cytometry Antibodies				
Target	Conjugation	Host	Source	Dilution	Target	Conjugation	Source	Clone	Dilution
dsRED	none	rabbit	Clontech (632496)	1:500	CD3	PerCP/Cy5.5	Biolegend	17A2	1:100
RFP	biotin	rabbit	Abcam (ab34771)	1:500	CD19	PerCP/Cy5.5	BD Biosciences	1D3	1:100
MUC2	none	rabbit	Santa Cruz (sc-15334)	1:100	CD11B	PerCP/Cy5.5	Biolegend	M1/70	1:300
LYS	none	rabbit	Dako (2017-06)	1:1000	CD5	PerCP/Cy5.5	eBiosciences	53-7.3	1:1000
CHGA	none	rabbit	Immunostar (20085)	1:250	NK1.1	PerCP/Cy5.5	eBiosciences	PK136	1:100
DCLK1	none	rabbit	Abcam (ab31704)	1:1000	EPCAM	PerCP/Cy5.5	Biolegend	G8.8	1:300
PTGS1	none	goat	Santa Cruz (sc-1754)	1:100	EPCAM	AF488	Biolegend	G8.8	1:300
GFI1B	none	goat	Santa Cruz (sc-8559)	1:100	GATA3	AF488	eBiosciences	TWAJ	2.5 μ l/test
EPCAM	AF488	rat	Biolegend (G8.8)	1:250	EPCAM	APC	Biolegend	G8.8	1:300
KI67	AF488	rat	Biolegend (16A8)	1:100	CD4	APC	BD Biosciences	RM4-5	1:100
rabbit IgG	AF488	goat	Life Technologies F(ab') ₂	1:1000	CD45	BV605	Biolegend	30-F11	1:100
rabbit IgG	AF555	goat	Life Technologies F(ab') ₂	1:1000	human CD4	PE	eBiosciences	RPA-T4	5 μ l/test
rabbit IgG	AF488	chicken	Life Technologies F(ab') ₂	1:1000	DCLK1	none	Abcam		1:1000
goat IgG	AF555	donkey	Life Technologies F(ab') ₂	1:1000	rabbit IgG	AF488	Life Technologies		1:2000

c		
qRT-PCR Primers		
Target	Forward Primer (5' -> 3')	Reverse Primer (5' -> 3')
<i>Dclk1</i>	CAAGCCAGCCATGTCGTTC	TTCCCTTTGAAGTAGCGGTAC
<i>Chga</i>	ATCCTCTCTATCCTGCGACAC	GGGCTCTGGTTCTCAAACACT
<i>Chat</i>	GGCCATTGTGAAGCGGTTTG	GCCAGGCGGTTGTTTAGATACA
<i>Trpm5</i>	TATGGCTTGTGGCCTATGGT	ACCAGCAGGAGAATGACCAG
<i>Muc2</i>	ATGCCACCTCCTCAAAGAC	GTAGTTTCCGTTGGAACAGTGAA
<i>Gnat3</i>	TAGGAGCCGAGAGGACCAAG	GCTGGTATTCAGATGCCCTTTC
<i>Lyz1</i>	GAGACCGAAGCACCGACTATG	CGGTTTTGACATTGTGTTCGC
<i>Lyz2</i>	ATGGAATGGCTGGCTACTATGG	ACCAGTATCGGCTATTGATCTGA
<i>Plgs1</i>	ATGAGTCGAAGGAGTCTCTCG	GCACGGATAGTAACAACAGGGA
<i>Gfi1b</i>	ATGCCACGGTCCTTTCTAGTG	GGAAGGCTCTGGTTCAGCAA
<i>Il25</i>	ACAGGGACTTGAATCGGGTC	TGGTAAAGTGGGACGGAGTTG
<i>Tslp</i>	ACGGATGGGGCTAACTTACAA	AGTCCTCGATTGCTCGAACT
<i>Il33</i>	GCTGCGTCTGTTGACACATTGAG	GGTCTTGCTCTTGGTCTTTTCCAG
<i>Rps17</i>	CGCCATTATCCCAGCAAG	TGTCGGATCCACCTCAATG

Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites

François Gerbe^{1,2,3}, Emmanuelle Sidot^{1,2,3}, Danielle J. Smyth^{4†}, Makoto Ohmoto⁵, Ichiro Matsumoto⁵, Valérie Dardalhon^{3,6}, Pierre Cesses^{1,2,3}, Laure Garnier^{1,2,3}, Marie Pouzolles^{3,6}, Bénédicte Brulin^{1,2,3}, Marco Bruschi^{1,2,3}, Yvonne Harcus⁴, Valérie S. Zimmermann^{3,6}, Naomi Taylor^{3,6}, Rick M. Maizels^{4†} & Philippe Jay^{1,2,3}

Helminth parasitic infections are a major global health and social burden¹. The host defence against helminths such as *Nippostrongylus brasiliensis* is orchestrated by type 2 cell-mediated immunity². Induction of type 2 cytokines, including interleukins (IL) IL-4 and IL-13, induce goblet cell hyperplasia with mucus production, ultimately resulting in worm expulsion^{3,4}. However, the mechanisms underlying the initiation of type 2 responses remain incompletely understood. Here we show that tuft cells, a rare epithelial cell type in the steady-state intestinal epithelium⁵, are responsible for initiating type 2 responses to parasites by a cytokine-mediated cellular relay. Tuft cells have a Th2-related gene expression signature⁶ and we demonstrate that they undergo a rapid and extensive IL-4R α -dependent amplification following infection with helminth parasites, owing to direct differentiation of epithelial crypt progenitor cells. We find that the *Pou2f3* gene is essential for tuft cell specification. *Pou2f3*^{-/-} mice lack intestinal tuft cells and have defective mucosal type 2 responses to helminth infection; goblet cell hyperplasia is abrogated and worm expulsion is compromised. Notably, IL-4R α signalling is sufficient to induce expansion of the tuft cell lineage, and ectopic stimulation of this signalling cascade obviates the need for tuft cells in the epithelial cell remodelling of the intestine. Moreover, tuft cells secrete IL-25, thereby regulating type 2 immune responses. Our data reveal a novel function of intestinal epithelial tuft cells and demonstrate a cellular relay required for initiating mucosal type 2 immunity to helminth infection.

Experimental subcutaneous infection of mice with *N. brasiliensis* (*Nb*) stage 3 larvae induces a typical type-2 response that involves a remodelling of epithelial cell populations, with goblet cell hyperplasia visible as soon as 5 days post-infection^{3,4}. *Nb* L3 larvae first migrate from their injection site to the lungs, where they moult to the L4 stage, are coughed up, and swallowed to reach the intestines (day 2 post infection) where they mature and lay eggs (starting 5 days post-infection). *Nb* induces a rapid and robust type 2 response, resulting in worm expulsion by 6–8 days post infection.

While the doublecortin-like kinase 1 (Dclk1)-expressing tuft cells represent only 0.4% of intestinal epithelial cells in naive mice⁵, we found that *Nb* infection resulted in a 8.5-fold expansion in tuft cells (Fig. 1a, b), first detected by 5 days post-infection in intestinal crypts, where proliferative epithelial progenitor cells reside, and also in the villi by 7 days post infection (Fig. 1c, Extended Data Fig. 1a). The kinetics of tuft cell expansion was equivalent to that of goblet cells (Fig. 1d, Extended Data Fig. 1b). Neo-differentiated tuft cells were indistinguishable from tuft cells present in naive mice, as evaluated by expression of established tuft cell markers, including Dclk1, Sry-related transcription factor 9 (Sox9), and phospholipase C gamma 2 (Plc γ 2) (Extended Data Fig. 1c)^{6–8}. All tuft cells, characterized by Dclk1 and growth factor

independent 1b (Gfi1b)⁸ expression also co-expressed the Pou domain, class 2, transcription factor 3 (Pou2f3) (Fig. 2a). In addition, rare (<3%, $n=400$ cells counted) Pou2f3⁺;Dclk1^{low} or Pou2f3⁺;Dclk1⁻ cells

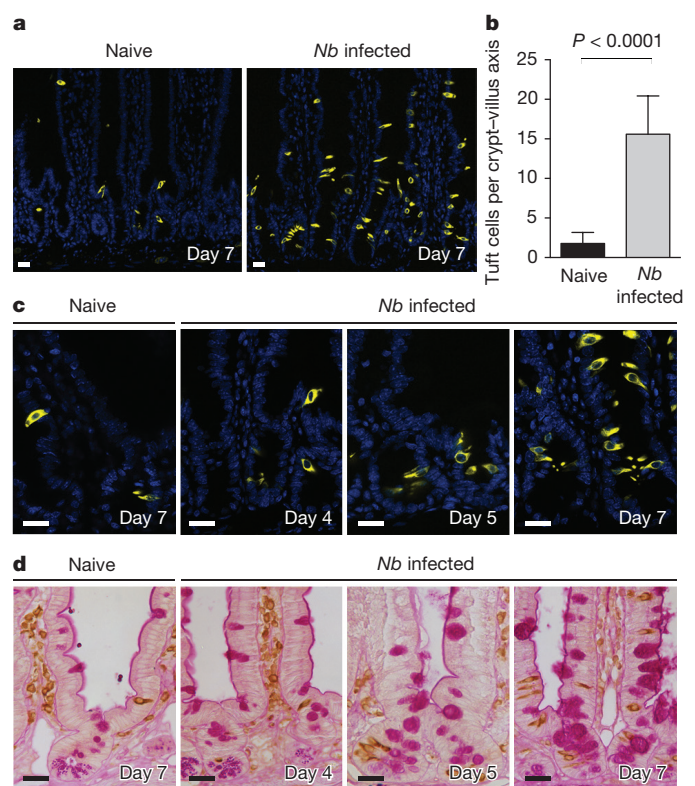


Figure 1 | Rapid amplification of the tuft cell lineage following infection with *Nb*. **a**, Presence of tuft cells in the intestinal epithelia of naive and *Nb*-infected mice 7 days post infection, visualized by expression of the Dclk1 marker. **b**, 8.7-fold increase of tuft cell numbers (1.8 ± 1.4 to 15.6 ± 4.8 per crypt-villus axis) in *Nb*-infected mice compared to naive mice, 7 days post infection. ($n=50$ crypt-villus units per mouse; 3 mice per condition). Data are shown as means \pm s.d. ($P < 0.0001$, two-tailed Student's *t*-test with Welch's correction). **c**, Changes in the Dclk1-expressing tuft cell population in intestinal crypts are presented at the indicated time points post infection. Quantification is shown in Extended Data Fig. 1a. **d**, Corresponding goblet cell hyperplasia associated with numerous and larger mucus vacuoles, detected by periodic acid-Schiff (PAS) staining. Dclk1 cells are also visualized in brown. Quantification is shown in Extended Data Fig. 1b. Scale bars, 20 μ m. All panels show representative pictures of experiments replicated 3 times in 3 mice per condition.

¹CNRS, UMR-5203, Institut de Génétique Fonctionnelle, F-34094 Montpellier, France. ²INSERM, U1191, F-34094 Montpellier, France. ³Université de Montpellier, F-34000 Montpellier, France.

⁴Institute of Immunology and Infection Research, School of Biological Sciences and Centre for Immunity, Infection and Evolution, University of Edinburgh, Edinburgh EH9 3JT, UK. ⁵Monell Chemical Senses Center, 3500 Market Street, Philadelphia, Pennsylvania 19104, USA. ⁶Institut de Génétique Moléculaire de Montpellier, CNRS, UMR5535, F-34293 Montpellier, France.

[†]Present address: Wellcome Trust Centre for Molecular Parasitology, Institute for Infection, Immunity and Inflammation, University of Glasgow, Glasgow G12 8TA, UK.

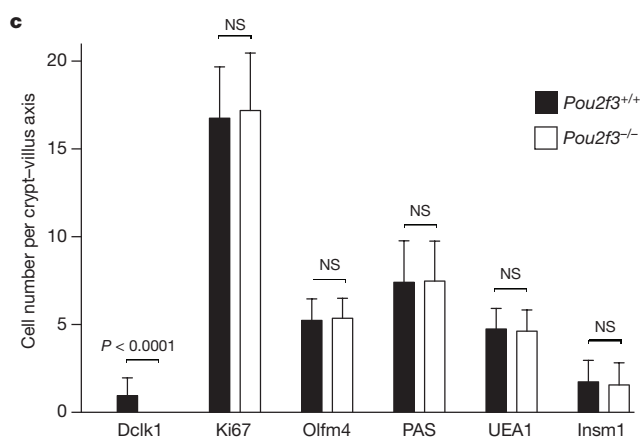
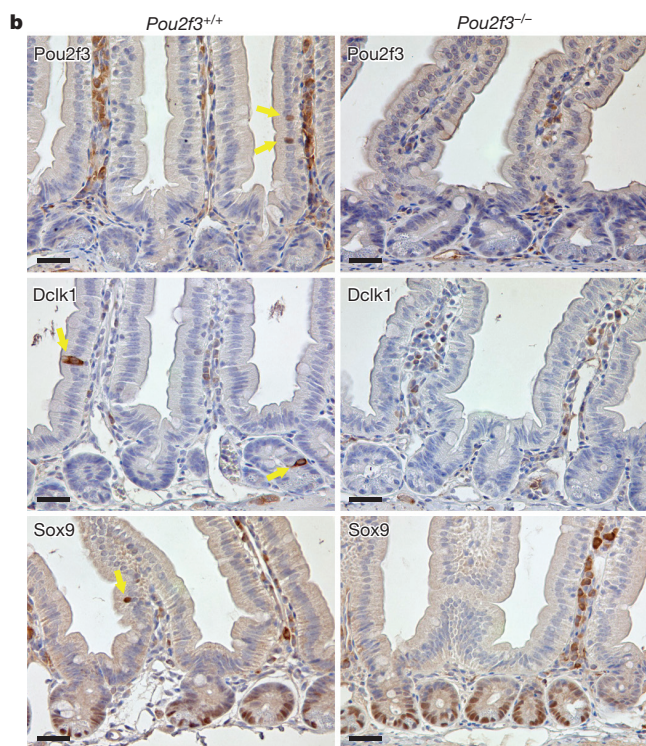
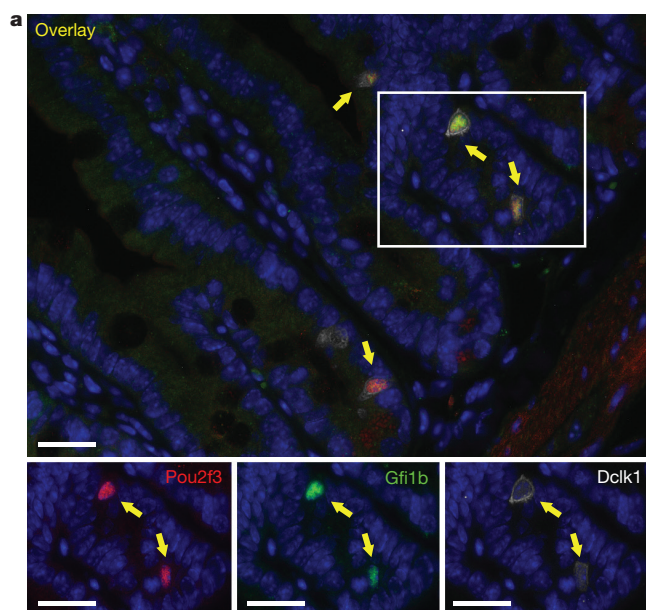


Figure 2 | Absence of tuft cells in the intestinal epithelium of *Pou2f3*^{-/-} mice. **a**, *Pou2f3* is expressed specifically in tuft cells of the intestinal epithelium as determined by co-staining for *Pou2f3* and established markers of tuft cells such as *Dclk1* and *Gfi1b*. **b**, *Pou2f3* deletion results in the absence of tuft cells as monitored by staining intestinal epithelium from *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice with *Pou2f3*-, *Dclk1*- and *Sox9*-specific antibodies. **a**, **b**, Three mice per genotype were used for staining experiments. Scale bars, 20 μ m. **c**, *Pou2f3* deficiency does not affect the proliferation zone ($P = 0.22$), stem cell compartment ($P = 0.66$), enterocyte (not counted), goblet ($P = 0.83$), Paneth ($P = 0.60$) or enteroendocrine ($P = 0.23$) cell lineages as monitored by Ki67, Olfm4, alkaline phosphatase, PAS staining, UEA1 lectin, and *Insm1*, respectively. ($n = 50$ crypt-villus units per mouse; 3 *Pou2f3*^{-/-} and 3 wild-type mice). Data are shown as means \pm s.d. A two-tailed Student's *t*-test was used. Pictures show representative experiments replicated 3 times.

were found at the base of crypts, probably representing early differentiating tuft cells since villus *Pou2f3*⁺ cells always co-express *Gfi1b* and *Dclk1*. Following infection, the percentage of proliferating tuft cells in crypts increased from $13 \pm 5.6\%$ to $24 \pm 14.9\%$ ($P = 0.035$), indicating that cell proliferation contributes to the amplification of the tuft lineage during type 2 responses (Extended Data Fig. 1d, e). Examination of the location of tuft cells present in *Nb*-infected mice revealed that some tuft cells differentiate close to the stem cell zone (Extended Data Fig. 1d), suggesting that biased differentiation from the recently described *Lgr5*⁺ slowly cycling early secretory progenitors⁹ and *Dll1*⁺ secretory progenitors¹⁰ also contributes to tuft cell lineage amplification. The increase in tuft cells was not due to a non-specific amplification of all secretory cell lineages as the number of enteroendocrine cells expressing the insulinoma-associated 1 (*Insm1*) marker¹¹, another secretory lineage of the intestinal epithelium, was significantly ($P = 0.008$) reduced (Extended Data Fig. 1f, g).

To determine whether the increase in the tuft cell population following infection with *Nb* was specific to C57BL/6 mice, we infected BALB/c mice and also observed a significant increase in tuft cell numbers (14-fold, $P < 0.0001$; Extended Data Fig. 2a, b). Moreover, this response seems to be a common adaptation to helminth infection in general, as infection of C57BL/6 and BALB/c mice strains with *Heligmosomoides polygyrus*¹² also resulted in a significant increase in tuft cell numbers (6.1- and 8.3-fold, respectively, $P < 0.0001$; Extended Data Fig. 2c, d). Tuft cell hyperplasia following *Nb* infection also occurred in *Rag*^{-/-} mice (10-fold; $P < 0.0001$) and therefore does not require functional adaptive immunity (Extended Data Fig. 2e, f).

Epithelial remodelling following helminth infection includes goblet cell hyperplasia and changes in mucus composition, associated with protective type 2 immunity^{13,14}. To investigate the role of tuft cells in this process, we identified and characterized a tuft-cell-deficient mouse line. Mice deficient for the *Pou2f3* transcription factor lack all *Pou2f3*-expressing taste receptor cells including sweet, umami and bitter taste cells¹⁵, as well as *Trpm5*-expressing chemosensory cells in the nasal cavity¹⁶ and olfactory epithelium¹⁷. Analysis of *Pou2f3*-deficient mice revealed a unique phenotype in the intestinal epithelium, with a complete absence of tuft cells as assessed by the absence of *Pou2f3*, *Dclk1* and *Sox9* expression outside the crypt compartment (Fig. 2b). The stem cell compartment, proliferation zone, and differentiation of enterocytes, goblet, enteroendocrine and Paneth cells were not affected (Fig. 2c and Extended Data Fig. 3). Furthermore, the distribution of immune cells in lymph nodes, mesenteric lymph nodes, spleen and lamina propria of *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice was equivalent (Extended Data Fig. 4) and lymphocytes were capable of responding to immune stimulation (Extended Data Fig. 5). Notably, type 2 innate lymphoid cells (ILC2s), a lineage that plays a critical role in secreting type 2 cytokines in response to helminth infection^{18,19}, were present in both the mesenteric lymph nodes and lamina propria of *Pou2f3*^{-/-} mice, at levels that were not significantly different from wild-type mice. (Extended Data Fig. 6a–c). Therefore, the absence of

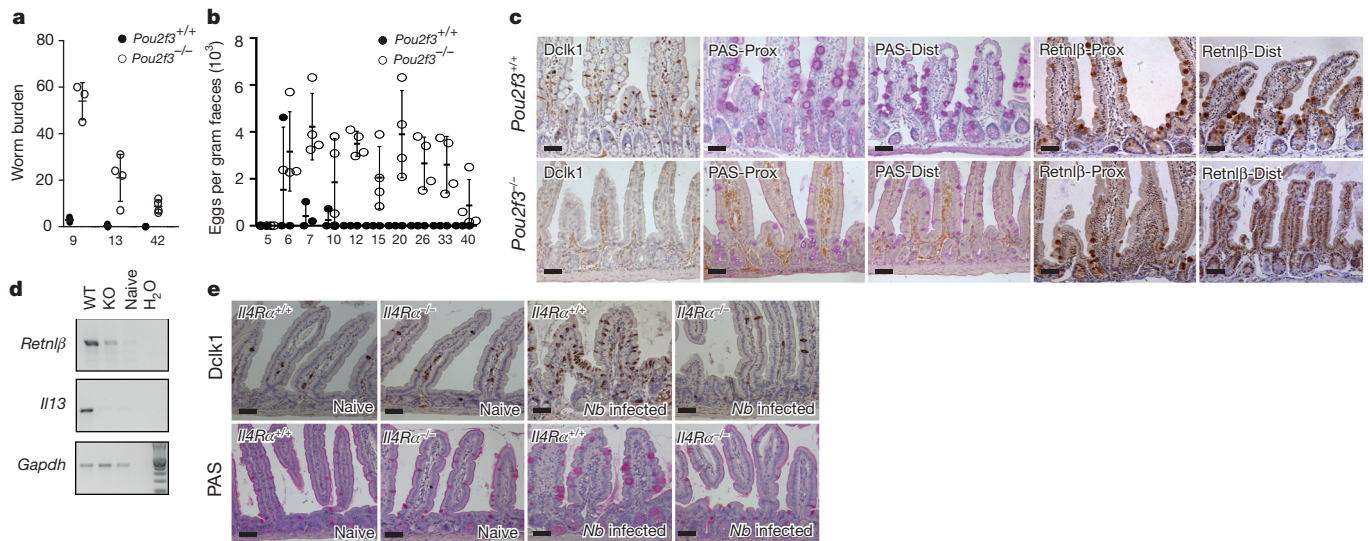


Figure 3 | Impaired type 2 responses in tuft cell-deficient mice.

a, Live adult worm counts in the small intestines of wild-type and *Pou2f3*^{-/-} mice at days 9, 13 and 42 post infection with *Nb* (*n* = 3 wild-type mice and 4 *Pou2f3*^{-/-} mice for each time point except for day 9 where *n* = 3 *Pou2f3*^{-/-} mice). **b**, Kinetic of *Nb* infection in 3 wild-type and 4 *Pou2f3*^{-/-} mice, assessed by faecal eggs count. **a**, **b**, Each circle represents an individual mouse. The *x* axis indicates time (days) post-infection. Average values ± s.d. are shown. **c**, Immunohistochemistry illustrating the proximal and distal small intestinal epithelium of infected wild-type and *Pou2f3*^{-/-} mice 7 days after infection (*n* = 3 mice per genotype).

Pou2f3 does not affect global immunity or intestinal epithelium formation. Rather, *Pou2f3* represents the first identified transcription factor that is specifically required to specify the tuft cell lineage in the intestinal epithelium, analogously to Sox9 for Paneth^{20,21} and Ngn3 for enteroendocrine²² cell lineages. Thus, *Pou2f3*^{-/-} mice represent a powerful model to study the function of tuft cells.

Pou2f3^{+/+} and *Pou2f3*^{-/-} mice were infected with *Nb* and analysed at several time points. In *Pou2f3*^{+/+} mice, only few worms were found after 9 days and expulsion was nearly complete after 13 days. In sharp contrast, numerous worms were found in *Pou2f3*^{-/-} mice up to 42 days post infection (Fig. 3a, b), not only in the proximal part of the small intestine, their normal site of attachment²³, but also in more distal locations. Together, these data strongly suggest that a compromised type-2 response is responsible for prolonged worm survival in *Pou2f3*^{-/-} tuft-cell-deficient mice.

To understand the mechanisms underlying the delayed worm expulsion in *Pou2f3*-deficient mice, we analysed the type-2 response-dependent remodelling of the intestinal epithelium 7 days after infection, a time point at which adult worms were detected in all infected animals. In *Pou2f3*^{+/+} mice, the intestinal epithelium displayed extensive and generalized goblet cell hyperplasia, with large mucus vacuoles, and tuft cell hyperplasia (Fig. 3c). Expectedly, *Pou2f3*^{-/-} mice completely lacked tuft cells and, in contrast to *Pou2f3*^{+/+} mice, were devoid of overt goblet cell hyperplasia, with focal and moderate hyperplasia limited to the most proximal small intestine, and lower goblet cell numbers than wild-type mice (Fig. 3c, Extended Data Fig. 7a, and Supplementary Information 1 and 2). Therefore, tuft-cell-deficient mice have a delayed type 2 response, with deficient mucosal goblet cell hyperplasia and delayed control of *Nb* infection.

The goblet cell-produced Resistin-like beta (Retnlβ) molecule, strongly induced by type 2 cytokines, has direct anti-helminth activity that facilitates expulsion^{3,24}. We compared expression of Retnlβ in wild-type and *Pou2f3*^{-/-} mice 7 days after *Nb* infection, when worm expulsion had started in wild-type mice. Retnlβ was strongly expressed in hyperplastic goblet cells in *Pou2f3*^{+/+} mice, but was only weakly expressed in *Pou2f3*^{-/-} mice (Fig. 3c, d and Extended Data Fig. 7a).

Dclk1 and PAS stainings, respectively, reveal tuft and goblet cells, as well as Retnlβ production. **d**, Quantification of IL-13 and Retnlβ in the intestinal mucosa of naive, and *Nb*-infected *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice by RT-PCR, 7 days after infection. Representative gels are shown with relative Gapdh expression presented as an internal control. **e**, Histological analysis showing tuft (Dclk1 staining) and goblet (PAS staining) cells in naive and *Nb*-infected *Il4Ra*^{+/+} and *Il4Ra*^{-/-} mice 7 days post infection (*n* = 3 mice per genotype). Scale bars, 20 μm. All panels show representative pictures of experiments replicated 3 times.

Moreover, while IL-4 levels were equivalent in mucosal tissue of *Nb*-infected *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice, IL-13 levels were markedly decreased in the latter (Fig. 3d). As both IL-4 and IL-13 type 2 cytokines are known to regulate Retnlβ expression³, and IL-4 is dispensable during type 2 responses to *Nb*²⁵, our data strongly suggest that defective IL-13 production is responsible for the decreased Retnlβ expression in *Nb*-infected *Pou2f3*^{-/-} mice. Thus, we identify a defective IL-13/Retnlβ axis in tuft-cell-deficient mice with impaired worm expulsion.

We next studied the link between tuft cells and type-2-mediated mucosal adaptation following *Nb* infection. IL-4Rα signalling is essential for both goblet cell hyperplasia and type 2 immune responses occurring upon helminth infection, and deletion of the *Il4ra* gene abrogates *Nb* expulsion^{23,26}. Importantly, the *Nb*-induced tuft cell hyperplasia occurring in wild-type mice 7 days post infection was absent in *Il4ra*^{-/-} mice, as was goblet cell hyperplasia (Fig. 3e, Extended Data Fig. 7b). This demonstrates the critical role of IL-4Rα signalling in the expansion of the tuft cell population following helminth infection.

We then examined whether IL-4Rα signalling is sufficient to trigger tuft cell lineage hyperplasia by injecting naive C57BL/6 mice with recombinant murine IL-4 and/or IL-13 (rIL-4/rIL-13) for 5 days and assessing the histology of the intestinal epithelium. rIL-4/rIL-13 injection induced goblet cell hyperplasia together with tuft cell expansion (Extended Data Fig. 7c). Importantly, treatment of *Pou2f3*^{-/-} mice with rIL-4/rIL-13 also resulted in goblet as well as Paneth cell hyperplasia, indicating a function of tuft cells upstream of IL-4/IL-13 (Extended Data Fig. 7c, d). Moreover, ectopic IL-4/IL-13 induced Retnlβ expression in goblet cells, independently of the *Pou2f3* genotype. Retnlβ expression was found predominantly in crypts and was therefore delayed compared to the onset of goblet cell hyperplasia (Extended Data Fig. 7c), and quantitatively lower than in an infectious context (Fig. 3c). Thus, IL-4Rα signalling is sufficient to induce an expansion of the tuft cell lineage. Furthermore, ectopic stimulation of this signalling cascade obviates the need for tuft cells in the epithelial cell remodelling of the intestine, including induction of Retnlβ expression by hyperplastic goblet cells.

To determine whether the IL-4/IL-13-induced goblet cell hyperplasia was epithelial-cell-autonomous, we used an *ex vivo* organoid culture

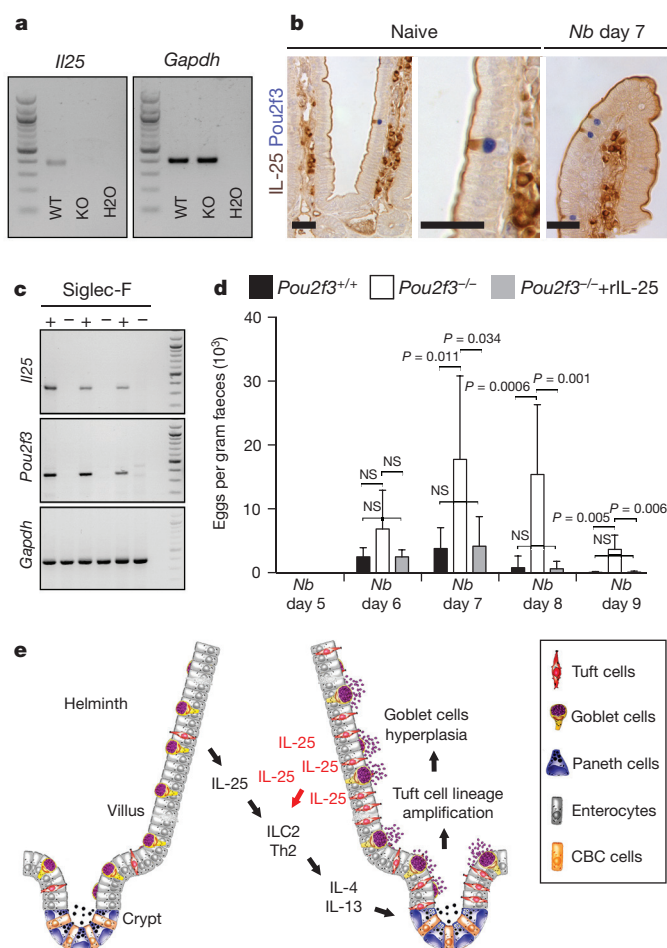


Figure 4 | Tuft cells express IL-25, and rIL-25 is sufficient to initiate type 2 mucosal responses in the absence of tuft cells. **a**, Analysis of *Il25* mRNA expression in *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice infected with *Nb*, 9 days post-infection, by RT-PCR. *Gapdh* expression is presented as an internal control. **b**, Immunohistochemistry showing IL-25 expression in naive and *Nb*-infected wild type mice. Blue staining, nuclear *Pou2f3* expression revealed with NBT/BCIP. Brown staining, IL-25 expression revealed with DAB ($n = 3$ naive and 3 infected mice). Scale bars, 20 μ m. **c**, PCR with reverse transcription (RT-PCR) showing predominant *Il25* and *Pou2f3* mRNA expression in the FACS-enriched tuft cells fractions (+) and the other epithelial cells (-), obtained from 3 independent mice. *Gapdh* is shown as an internal control. **d**, Rescue of the *Pou2f3* deficiency by treatment with exogenous rIL-25, as assessed by egg counts during a time course of infection with *Nb* ($n = 7$ mice for the wild-type and *Pou2f3*^{-/-} mice, and $n = 6$ for the rIL-25-treated *Pou2f3*^{-/-} mice). Average \pm s.d. are presented, as well as exact P values when < 0.05 (two-tailed Mann-Whitney U -test). **e**, Scheme illustrating the function of tuft cells in initiating type 2 responses following infection with intestinal helminths. Left, normal epithelium undergoing infection with a helminth. Right, tuft cell-dependent epithelial remodelling during type 2 responses. All panels show representative pictures of experiments replicated 3 times.

system²⁷ that allows physiological responses of an isolated intestinal epithelium to be analysed in the absence of stromal cues. As expected, tuft cells were absent in *Pou2f3*^{-/-} organoid cultures (Extended Data Fig. 8a). Moreover, in wild-type organoids, the tuft cell population increased as early as 48 h following addition of rIL-4/rIL-13 (Extended Data Fig. 8a, b). Treatment with rIL-4 or rIL-13 alone yielded identical results to the rIL-4/rIL-13 mixture (Extended Data Fig. 8c). Treatment of *Pou2f3*^{-/-} organoids with rIL-4/rIL-13 also triggered goblet cell hyperplasia equivalent to that detected in *Pou2f3*^{+/+} organoids, as indicated by Retnl β expression (Extended Data Fig. 8d), revealing the critical role of type 2 cytokine signalling downstream of the tuft

cell lineage. Furthermore, these data demonstrate that the intestinal epithelial response to IL-4/IL-13 is epithelium-autonomous and does not require additional stromal signals. Together, our data identify a novel function of tuft cells in initiating the mucosal type 2 responses with a positive feedback loop through IL-13-producing immune cells that, in turn, amplify the tuft cell lineage.

Finally, we investigated the physiological function of the tuft cell hyperplasia, fully established by 7 days post-infection when worm expulsion starts. IL-25 is an alarmin molecule produced by an as yet unidentified intestinal epithelial cell type, capable of initiating type 2 responses by stimulating ILC2s to produce IL-4 and IL-13^{18,19}. We thus analysed *Il25* messenger RNA expression in *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice infected with *Nb*. Nine days after infection, *Il25* expression was higher in the intestinal mucosa of *Pou2f3*^{+/+} mice than in tuft-cell-deficient *Pou2f3*^{-/-} mice (Fig. 4a). Moreover, IL-25 protein expression was restricted to tuft cells in naive mice (Fig. 4b and Extended Data Fig. 9a) and consistent with these data, *Il25* mRNA was only detected in the FACS-enriched tuft cell fraction of the intestinal epithelium (Fig. 4c and Extended Data Fig. 9b). Following *Nb* infection, IL-25 expression remained restricted to tuft cells (Fig. 4b). Concomitant with tuft cell hyperplasia, epithelial IL-25 expression peaks 9 days after infection with *Nb*, at the time of worm expulsion, for which it is required²⁸. In accord with a critical role for IL-25-secreting tuft cells in the expansion of ILC2s, we found that the percentage of Lin⁻CD127⁺Gata3⁺KLRG1⁺ ILC2s was not significantly augmented by *Nb* infection of *Pou2f3*^{-/-} mice, but was significantly augmented in wild-type mice. Indeed, tuft cells were required for the global induction of an adaptive immune response as helminth infection induced an approximately 2.5-fold expansion of both ILC2 and Th2 subsets in mesenteric lymph nodes, whereas these subsets remained unchanged in the infected *Pou2f3*^{-/-} mice ($P = 0.02$, $P = 0.0005$, respectively; Extended Data Fig. 6d–f). It is likely that these immune defects are directly due to the paucity of IL-25 as treatment of *Nb*-infected *Pou2f3*^{-/-} mice with rIL-25 almost completely compensated for the absence of tuft cells, promoting an efficient worm expulsion (Fig. 4d). IL-25 thus provides a mechanistic link between tuft cells, promotion of type 2 responses and worm expulsion.

Taken together, our data reveal a critical function of tuft cells in initiating mucosal type 2 responses following infection with helminths through IL-25 secretion. In the absence of tuft cells, IL-25 and IL-13 expression remain low, and type 2 mucosal responses and worm expulsion are delayed. Our study demonstrates a requirement for tuft cells upstream of IL-4/IL-13, with these cytokines driving tuft cell hyperplasia, thereby amplifying a feed-forward loop to orchestrate a rapid and effective anti-helminth immunity (Fig. 4e).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 August; accepted 10 December 2015.

- Hotez, P. J. *et al.* Helminth infections: the great neglected tropical diseases. *J. Clin. Invest.* **118**, 1311–1321 (2008).
- Allen, J. E. & Maizels, R. M. Diversity and dialogue in immunity to helminths. *Nature Rev. Immunol.* **11**, 375–388 (2011).
- Herbert, D. R. *et al.* Intestinal epithelial cell secretion of RELM-beta protects against gastrointestinal worm infection. *J. Exp. Med.* **206**, 2947–2957 (2009).
- McKenzie, G. J., Bancroft, A., Grecis, R. K. & McKenzie, A. N. A distinct role for interleukin-13 in Th2-cell-mediated immune responses. *Curr. Biol.* **8**, 339–342 (1998).
- Gerbe, F. *et al.* Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *J. Cell Biol.* **192**, 767–780 (2011).
- Bezençon, C. *et al.* Murine intestinal cells expressing Trpm5 are mostly brush cells and express markers of neuronal and inflammatory cells. *J. Comp. Neurol.* **509**, 514–525 (2008).
- Gerbe, F., Brulin, B., Makrini, L., Legraverend, C. & Jay, P. DCAMKL-1 expression identifies tuft cells rather than stem cells in the adult mouse intestinal epithelium. *Gastroenterology* **137**, 2179–2180 (2009).

8. Bjerknes, M. *et al.* Origin of the brush cell lineage in the mouse intestinal epithelium. *Dev. Biol.* **362**, 194–218 (2012).
9. Buczaczk, S. J. *et al.* Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* **495**, 65–69 (2013).
10. van Es, J. H. *et al.* Dll1⁺ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biol.* **14**, 1099–1104 (2012).
11. Gierl, M. S., Karoulias, N., Wende, H., Strehle, M. & Birchmeier, C. The zinc-finger factor Insm1 (IA-1) is essential for the development of pancreatic β cells and intestinal endocrine cells. *Genes Dev.* **20**, 2465–2478 (2006).
12. Reynolds, L. A., Filbey, K. J. & Maizels, R. M. Immunity to the model intestinal helminth parasite *Heligmosomoides polygyrus*. *Semin. Immunopathol.* **34**, 829–846 (2012).
13. Ishikawa, N., Horii, Y. & Nawa, Y. Immune-mediated alteration of the terminal sugars of goblet cell mucins in the small intestine of *Nippostrongylus brasiliensis*-infected rats. *Immunology* **78**, 303–307 (1993).
14. Watanabe, N., Katakura, K., Kobayashi, A., Okumura, K. & Ovary, Z. Protective immunity and eosinophilia in IgE-deficient SJA/9 mice infected with *Nippostrongylus brasiliensis* and *Trichinella spiralis*. *Proc. Natl Acad. Sci. USA* **85**, 4460–4462 (1988).
15. Matsumoto, I., Ohmoto, M., Narukawa, M., Yoshihara, Y. & Abe, K. Skn-1a (Pou2f3) specifies taste receptor cell lineage. *Nature Neurosci.* **14**, 685–687 (2011).
16. Ohmoto, M. *et al.* Pou2f3/Skn-1a is necessary for the generation or differentiation of solitary chemosensory cells in the anterior nasal cavity. *Biosci. Biotechnol. Biochem.* **77**, 2154–2156 (2013).
17. Yamaguchi, T. *et al.* Skn-1a/Pou2f3 is required for the generation of Trpm5-expressing microvillous cells in the mouse main olfactory epithelium. *BMC Neurosci.* **15**, 13 (2014).
18. Moro, K. *et al.* Innate production of T_H2 cytokines by adipose tissue-associated c-Kit⁺Sca-1⁺ lymphoid cells. *Nature* **463**, 540–544 (2010).
19. Neill, D. R. *et al.* Nuocytes represent a new innate effector leukocyte that mediates type-2 immunity. *Nature* **464**, 1367–1370 (2010).
20. Bastide, P. *et al.* Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *J. Cell Biol.* **178**, 635–648 (2007).
21. Blache, P. *et al.* SOX9 is an intestine crypt transcription factor, is regulated by the Wnt pathway, and represses the *CDX2* and *MUC2* genes. *J. Cell Biol.* **166**, 37–47 (2004).
22. Mellitzer, G. *et al.* Loss of enteroendocrine cells in mice alters lipid absorption and glucose homeostasis and impairs postnatal survival. *J. Clin. Invest.* **120**, 1708–1721 (2010).
23. Camberis, M., Le Gros, G. & Urban, J. Jr. Animal model of *Nippostrongylus brasiliensis* and *Heligmosomoides polygyrus*. *Curr. Protoc. Immunol.* **Ch. 19**, Unit 19.12 (2003).
24. Artis, D. *et al.* RELM β /FIZZ2 is a goblet cell-specific immune-effector molecule in the gastrointestinal tract. *Proc. Natl Acad. Sci. USA* **101**, 13596–13600 (2004).
25. Lawrence, R. A., Gray, C. A., Osborne, J. & Maizels, R. M. *Nippostrongylus brasiliensis*: cytokine responses and nematode expulsion in normal and IL-4-deficient mice. *Exp. Parasitol.* **84**, 65–73 (1996).
26. Urban, J. F. Jr *et al.* IL-13, IL-4R α , and Stat6 are required for the expulsion of the gastrointestinal nematode parasite *Nippostrongylus brasiliensis*. *Immunity* **8**, 255–264 (1998).
27. Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures *in vitro* without a mesenchymal niche. *Nature* **459**, 262–265 (2009).
28. Zhao, A. *et al.* Critical role of IL-25 in nematode infection-induced alterations in intestinal function. *J. Immunol.* **185**, 6921–6929 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by Association pour la Recherche contre le Cancer (ARC to N.T. and SL220110603456 to P.J.), Agence Nationale de la Recherche (ANR-09-BLAN-0368-01 to P.J., ANR-PolarAttack to V.D. and ANR-14-CE14-0025-01 to P.J. and N.T.), Institut National du Cancer (INCa 2014-174 to P.J.), a CNRS-NIH International Laboratory grant from the CNRS (LIA-BAGEL) to N.T. and the Wellcome Trust (Ref. 106122 to R.M.M.). Part of the work was supported by institutional funds of Monell Chemical Senses Center to I.M.; E.S. and M.B. are supported by Ligue Nationale contre le Cancer, M.P. by the LabEx EpiGenMed and N.T. by Inserm. We are grateful to S. Gailhac and C. Mongellaz for their expertise and assistance in immune cell analyses, S. Cording, J. Di Santo and G. Eberl for their generosity and their expertise on ILCs, G. Petrazzo and R. Guédon for technical input, S. Fre and M. Huygues for technical advice, C. Legraverend for her expertise and to M. van de Wetering for reagents. We thank F. Gallardo, the PCEA and the IRD A2 facilities for maintenance of mouse colonies, the Montpellier RIO Imaging (MRI) facility, the Monell Histology and Cellular Localization Core (supported by funding from NIH Core Grant P30DC011735 (to R. F. Margolske, Monell Chemical Senses Center) for some of the histological analyses and Daniel Fisher for critical reading of the manuscript.

Author Contributions F.G. performed the majority of the experiments. E.S., D.J.S., B.B. and P.C. contributed to mouse studies, M.O. and I.M. to characterization of the Pou2f3-deficient mouse line, E.S., L.G. and M.B. to organoid experiments, V.D., M.P. and V.S.Z. to immune studies and Y.H. to parasite life cycle experiments. P.J. and F.G. conceived the study. P.J., F.G. and R.M.M. designed experiments with contributions from V.D., V.S.Z. and N.T.; P.J. wrote the manuscript with inputs from F.G. and N.T.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.J. (philippe.jay@igf.cnrs.fr).

METHODS

Animal strains. The *Pou2f3*-deficient mice (*Pou2f3^{tm1Abek}*) have been previously described¹⁵. *Il4ra*-deficient mice²⁹ were provided by M. Kopf (Basel Institute for Immunology, Switzerland). C57BL/6 and BALB/c mice were obtained from Charles River Laboratories. All the mice were maintained in an SPF animal facility and were naive before the experiments. All animal experiments were approved by the Institutional Animal Care and Use Committee of Monell Chemical Senses Center or by the French Agriculture and Forestry Ministry. Unless specified, all mice were on a C57BL/6 genetic background. Mice were analysed at 10 weeks of age, regardless of the sex. For comparisons of wild-type and KO mice, littermates were analysed. Three mice per condition were analysed in all experiments. No statistical methods were used to predetermine sample size, no criteria of exclusion were defined and the experiments were not randomized nor blinded to the investigator.

Immunophenotyping and flow cytometry analyses. Cells, isolated from peripheral lymph nodes, mesenteric lymph nodes, spleen and lamina propria were stained with Sytox blue or Live/dead fixable viability dye (Life Technologies) and eBioscience respectively together with the appropriate conjugated anti-CD3, CD45.2, CD62L, CD4, CD8, CD44, TCR $\gamma\delta$, CD19, NK1.1, Gr1, CD11b, CD11c, and Foxp3 antibodies (eBioscience or Becton Dickinson, San Diego, CA). For ILC2 staining, cells were stained with a lineage cocktail and CD45.2⁺ (clone 104) and lineage-negative CD45⁺ cells were assessed for expression of CD127 (clone SB14), KLRG1 (clone 2F1), Sca-1 (clone D7), CD25 (clone 7D4), and intracellular expression of Gata-3 (clone L50-823). Th2 cells were identified on the basis of Gata-3-expressing cells within the CD3⁺CD4⁺ subset.

IL-6, IL-12, TNF β , IL-10, MCP-1 and IFN- γ production was assessed in the culture supernatant of LPS/IL-4-activated splenocytes using a Cytometric Bead Array (CBA) Kit (BD Biosciences). To assess intracellular cytokine production, freshly isolated and anti-CD3/CD28 stimulated LN cells were activated with PMA (Sigma-Aldrich; 100 ng ml⁻¹)/ionomycin (Sigma-Aldrich; 1 μ g ml⁻¹) in the presence of brefeldin A (Sigma-Aldrich; 10 μ g ml⁻¹) for 4 h at 37°C, fixed, permeabilized and then stained with specific antibodies against IL-2 and IFN- γ .

Foxp3 staining was performed following fixation/permeabilization (eBioscience). Stained cells were assessed by flow cytometry (LSR Fortessa, Becton Dickinson, San Jose, CA) and data were analysed by FACSDiva (v.8.0, BD Biosciences) and FCAP Array Software (CBA analysis).

Ex vivo stimulations. LN cell activation was performed using plate-bound anti-CD3 (clone 2C11; 1 μ g ml⁻¹) and anti-CD28 (clone PV-1 1 μ g ml⁻¹) mAbs in RPMI 1640 media (Life Technologies) supplemented with 10% FCS, 2 mM glutamine and 1% penicillin/streptomycin. Exogenous IL-2 (100 U ml⁻¹) was added every other day starting at day 2 post-activation. Cell proliferation was monitored by labelling with CFSE (Life Technologies; 2.5 μ M) for 3 min at room temperature. Splenocytes were activated with LPS (20 μ g ml⁻¹) and IL4 (25 ng ml⁻¹). Supernatants were collected 40 h following activation.

Immunoglobulin detection. IgG detection in supernatants of LPS/IL-4-stimulated splenocytes was assayed by ELISA. Microtiter plates (Maxisorb, Nunc) were saturated overnight at 4°C with 100 μ l of anti-IgG2a, anti-IgG2b, anti-IgA antibodies or anti-IgG (Fab'2) resuspended in PBS (5 μ g ml⁻¹). Plates were washed 3 times with 0.1% Tween-containing PBS (PBST). Samples (1/2 dilution) were diluted in a final volume of 100 μ l per well of PBST-1%BSA and incubated for 2 h at RT. Following washes, peroxidase-conjugated anti-mouse anti-IgG2a, anti-IgG2b, anti-IgA (Serotech) or anti-IgG gamma-chain (SIGMA) antibodies were added in PBST-1%BSA (1:1,000 dilution; 100 μ l per well) and incubated for a 1 h at 37°C. Immunoglobulin levels were then revealed with o-phenylenediamine (Sigma; 4 μ g ml⁻¹) in 0.1 M Na citrate and 0.03% hydrogen peroxide. Absorbance was measured at 450 nm using an automated plate reader (InfiniteM200Pro, TECAN) after 5 min at room temperature.

Tuft cell sorting. Freshly isolated small intestines of BC57BL/6 mice were flushed with PBS and incised along their length. The tissue was then incubated in 30 mM EDTA (Sigma) in HBSS pH 7.4 (Life Technologies) on ice, and transferred in DMEM (Life Technologies) supplemented with 10% FBS (Sigma). Vigorous shaking yielded the epithelial fraction that was then incubated with 100 μ l of Dispase (BD Biosciences) in 10 ml of HBSS, supplemented with 100 μ l of DNase I at 2,000 Kunitz (Sigma). Single cell preparation obtained by filtration on a 30 μ m mesh was incubated with a phycoerythrin rat anti-mouse Siglec-F antibody for 30 min at 4°C (BD pharmingen 552126), washed with HBSS and resuspended in appropriate volume of HBSS pH 7.4 supplemented with 5% FBS before staining with 7-aminocytinomylin D (Life Technologies) to exclude dead cells. Siglec-F⁺ live cells were sorted using a FACSAria (Becton Dickinson), directly in RLT lysis buffer (Qiagen) for subsequent RNA extraction.

Parasite infections. *Pou2f3^{+/+}*, *Pou2f3^{-/-}*, *Il4ra^{+/+}* and *Il4ra^{-/-}*, C57BL/6 and BALB/c wild-type mice were used for *Nb* infection experiments. Mice were infected with 250 L3 infective *Nb* larvae by sub-cutaneous injection²³ or with 200

H. polygyrus L3 larvae by gavage. Infection parameters were monitored by enumeration of live adult worms in the small intestinal tissue by two different investigators blinded to the study groups.

Reagents. Recombinant murine IL-4 (214-14), recombinant murine IL-13 (210-13) were purchased from PeproTech, and recombinant murine IL-25 (1399) was from R&D Systems. For animal treatment, mice were injected intraperitoneally daily with a mixture of both interleukins or with rIL-25 (40 μ g per kg of body weight). For rescue experiments in *Pou2f3^{-/-}* mice, rIL-25 was injected from day 3 post infection.

Organoid culture. Organoid cultures were performed as previously described²⁷ using intestinal crypts from *Pou2f3^{+/+}* and *Pou2f3^{-/-}* mice. Organoid lines were passaged up to 10 times before experiments to ensure pure epithelial cultures. When indicated, cultures were stimulated with recombinant murine IL-4 (400 ng ml⁻¹), recombinant murine IL-13 (400 ng ml⁻¹) or an equimolar mixture of the two cytokines. For histological studies, organoids were washed twice in cold PBS to dissolve the Matrigel, fixed overnight in neutral-buffered formalin at 4°C and included in Histogel (Thermo Scientific) before paraffin embedding ($n = 3$ experiments from independent mice).

RNA extraction and PCR. Total RNA from intestinal organoids or snap-frozen intestinal tissues was isolated using TRIzol (Life Technologies) followed by precipitation with isopropanol. Further RNA purification was carried out on RNeasy columns (Qiagen, 74104) and DNase treatment. In the case of Siglec⁺ sorted tuft cells, extraction and DNase treatment were performed using RNeasy Micro KIT following the manufacturer's instructions. Reverse transcription was performed with 500 ng–2 μ g of purified RNA using Transcriptor First Strand cDNA synthesis KIT (Roche) according to the manufacturer's instructions. For qRT-PCR experiments, gene expression was quantified on the LightCycler 480 using LightCycler 480 SYBR Green I Master (Roche). The results from three independent organoid cultures were normalized to the expression level of *Gapdh* and *Hprt* and relative expression was obtained using the $\Delta\Delta C_t$ method. Primer sets for each gene are listed in Extended Data Table 1. PCR analyses were performed on an Eppendorf Mastercycler, using the primer sets listed in the Extended Data Table 1.

In situ hybridization. Single colorimetric and double fluorescent *in situ* hybridization analyses were carried out as described previously³⁰. Briefly, digoxigenin- and fluorescein-labelled antisense RNAs were synthesized and used as probes after fragmentation to about 150 bases under alkaline conditions. Small intestines were dissected from mice shortly after euthanasia and embedded in frozen O.C.T. compound (Sakura Finetech). Fresh-frozen sections were prepared using a cryostat (CM1900, Leica Microsystems), fixed with 4% paraformaldehyde, hybridized with probe(s), and then washed under stringent conditions. Hybridized probes were immunohistochemically detected using alkaline phosphatase-conjugated anti-digoxigenin antibody (Roche Diagnostics) and biotin-conjugated anti-fluorescein antibody (Vector Laboratories). Signals were developed using 4-nitro blue tetrazolium chloride/5-bromo-4-chloro-3-indolyl-phosphate as chromogenic substrates for single colorimetric analyses or the Tyramid Signal Amplification method and HNPP Fluorescent Detection Set (Roche Diagnostics) for double-fluorescent analyses. Stained and fluorescent images were obtained on a Nikon eclipse 80i microscope (Nikon Instruments Inc.) equipped with a DXM1200C digital camera (Nikon) and a Leica SP2 confocal scanning microscope (Leica), respectively. RNA probes generated were as follows: nucleotides 72–2363 of *Pou2f3* (GenBank accession number NM_011139), nucleotides 1–2228 of *Slc15a1* (GenBank accession number BC116248), nucleotides 1–3255 of *Muc2* (GenBank accession number BC034197), nucleotides 1–1102 of *Gcg* (GenBank accession number BC012975), nucleotides 1–584 of *Gip* (GenBank accession number BC104314), nucleotides 27–400 of *Defcr6* (GenBank accession number M33225), nucleotides 1–1628 of *Olfm4* (GenBank accession number BC141127), nucleotides 1–2750 of *Dcl1* (GenBank accession number BC050903), and nucleotides 1–2797 of *Ptgs1* (GenBank accession number BC005573).

Fluorescent and bright-field immunohistochemistry on paraffin-embedded tissue. Tissue dissection, fixation, and immunohistochemistry on thin sections of paraffin-embedded tissue were performed essentially as described previously⁵. Primary antibodies used in this study were as follows: anti-Sox9 (AB5535; Millipore), anti-Cox1 (sc-1754; Santa Cruz), anti-PCNA (sc-56; Santa Cruz), anti-Plc γ 2 (sc-5283, Santa Cruz), anti-Gfi1b (Sc-8559; Santa Cruz), anti-*Pou2f3* (sc-330, Santa Cruz and HPA019652, Prestige Antibodies), anti-*Dcl1* (ab31704; AbCam), anti-Ki67 (ab16667; AbCam), anti-Retn β (ABIN465494, Antibodies online), anti-IL-25 (mAb 1258; R&D Systems). Anti-Insm1 was a gift from C. Birmeier (Max-Delbrück-Center for Molecular Medicine; Berlin; Germany). Slides were washed twice times with 0.1% PBS-Tween (Sigma-Aldrich) before incubation with fluorescent secondary antibodies conjugated with either Alexa 488, cyanin-3, or cyanin-5 (Jackson ImmunoResearch Laboratories, Inc.) and Hoechst at 2 μ g ml⁻¹ (Sigma-Aldrich) in PBS–Triton X-100 0.1%

(Sigma-Aldrich). Stained slides were washed again in PBS before mounting with FluoroMount (Sigma-Aldrich). Methods used for bright-field immunohistochemistry were identical, except that Envision+ (Dako) was used as a secondary reagent. Signals were developed with DAB (Sigma-Aldrich) and a haematoxylin counterstain (DiaPath) was used. After dehydration, sections were mounted in Pertex (Histolab). All experiments were performed on formalin-fixed tissues and 10 mM sodium citrate (pH 6.4) treated slides, except for IL-25 staining where Carnoy's fixation and 10 mM Tris-EDTA (pH 9.0) treatment were used. Enterocytes-alkaline phosphatase activity was revealed with Fast-red substrate (Sigma-Aldrich). All stainings were repeated in 3 mice per group in 3 independent experiments and scored by three different investigators blinded to the study groups.

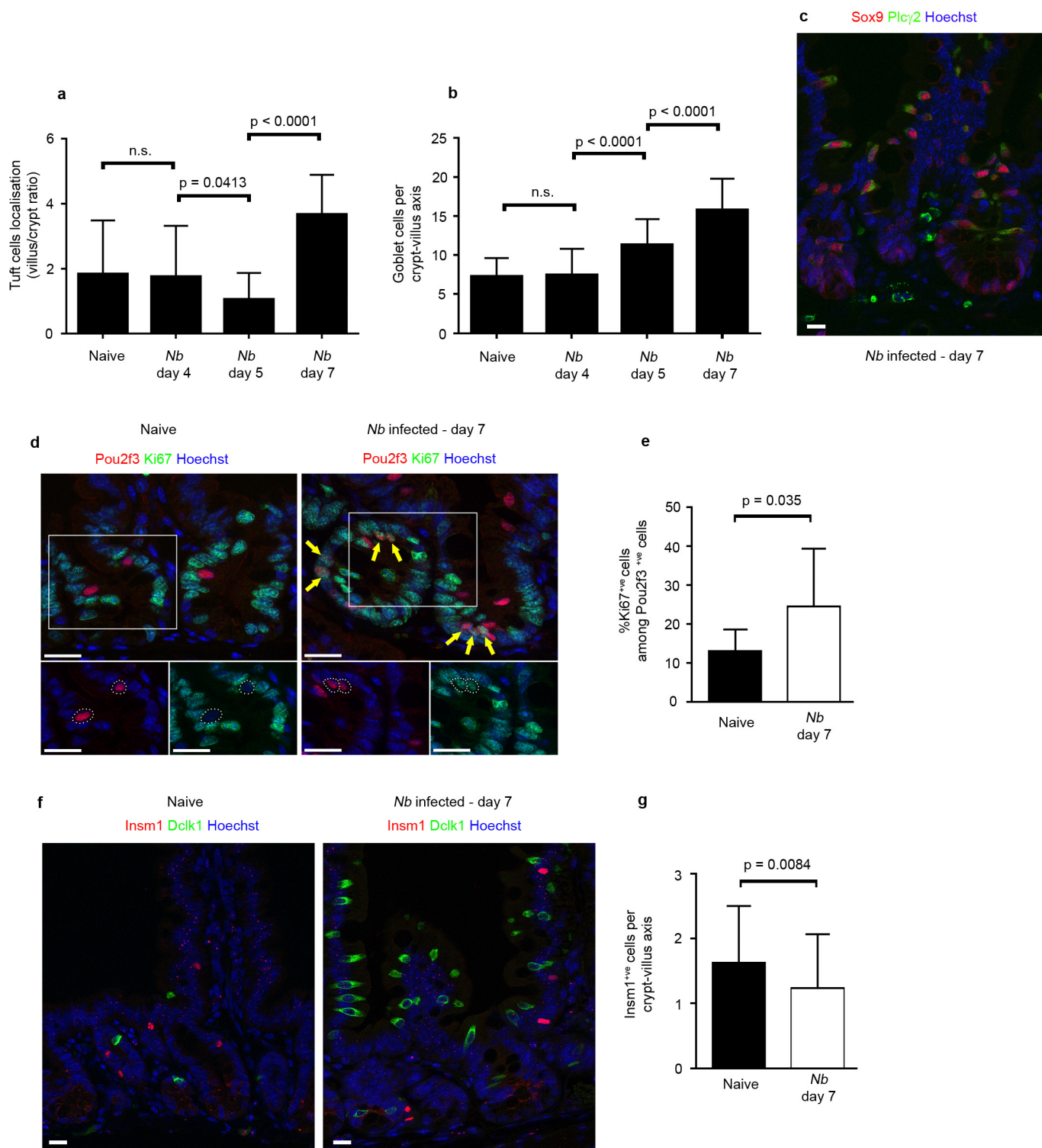
Microscopy and imaging. Fluorescent pictures were acquired at room temperature on an AxioImager Z1 microscope (Carl Zeiss, Inc.) equipped with a camera (AxioCam MRm; Carl Zeiss, Inc.), EC Plan Neofluar (5×, NA 0.16; 10×, NA 0.3; 20×, 0.5 NA; 100×, NA 1.3) and Plan Apochromat (40×, NA 0.95; 63×, NA 1.4) objectives, the Apotome Slider system equipped with an H1 transmission grid (Carl Zeiss, Inc.), and AxioVision software (Carl Zeiss, Inc.). Bright-field immunohistochemistry pictures were taken at room temperature on an Eclipse 80i microscope (Nikon) with Plan Fluor (10×, NA 0.3; 20×, NA 0.5; 40×, NA 0.75; and 60×, NA 0.5–1.25) lenses (Nikon) and a digital camera (Q-Imaging Retiga 2000R with a Q-Imaging RGB Slider). Pictures were captured with Q-Capture Pro software (Nikon). Post-treatment of pictures (level correction), annotations, and panel composition were performed using Photoshop software (Adobe).

Statistical analyses. The Prism software was used for descriptive statistical analyses. For infection monitoring, sample (n) was defined as the number of eggs per gram of faeces per mouse. As normal distribution assumption was not met, a two-tailed Mann–Whitney U -test was used to calculate the P value. For histological data quantification, sample (n) was defined as the number of cells per crypt–villus unit. Unless otherwise stated, 50 crypt–villus axes were counted per histological sections from 3 mice of each genotype or condition. According to the central limit theorem ($n > 30$), data comparison was achieved with a two-tailed Student's t -test. Welch's correction was applied to P -value calculation when homoscedasticity assumption was not met (determined with F -test).

Results are shown as histograms representing means as centre values and standard deviation as error bars for each genotype or conditions, except when $n < 5$, where individual data point are plotted.

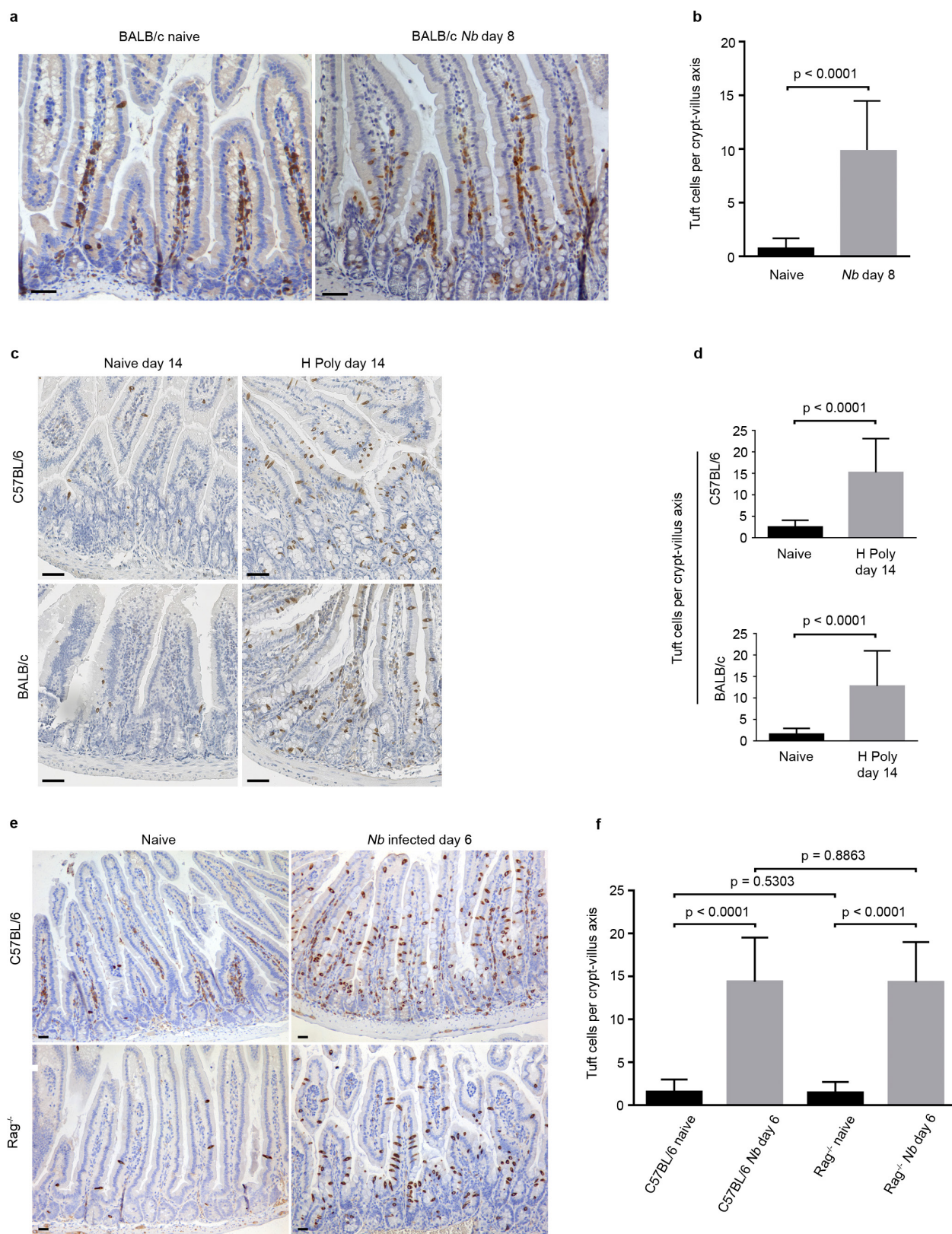
No statistical methods were used to predetermine sample size and the experiments were not randomized. Unless otherwise stated, the investigators were not blinded to allocation during experiments and outcome assessment.

29. Barner, M., Mohrs, M., Brombacher, F. & Kopf, M. Differences between IL-4R α -deficient and IL-4-deficient mice reveal a role for IL-13 in the regulation of Th2 responses. *Curr. Biol.* **8**, 669–672 (1998).
30. Ohmoto, M., Matsumoto, I., Yasuoka, A., Yoshihara, Y. & Abe, K. Genetic tracing of the gustatory and trigeminal neural pathways originating from T1R3-expressing taste receptor cells and solitary chemoreceptor cells. *Mol. Cell. Neurosci.* **38**, 505–517 (2008).



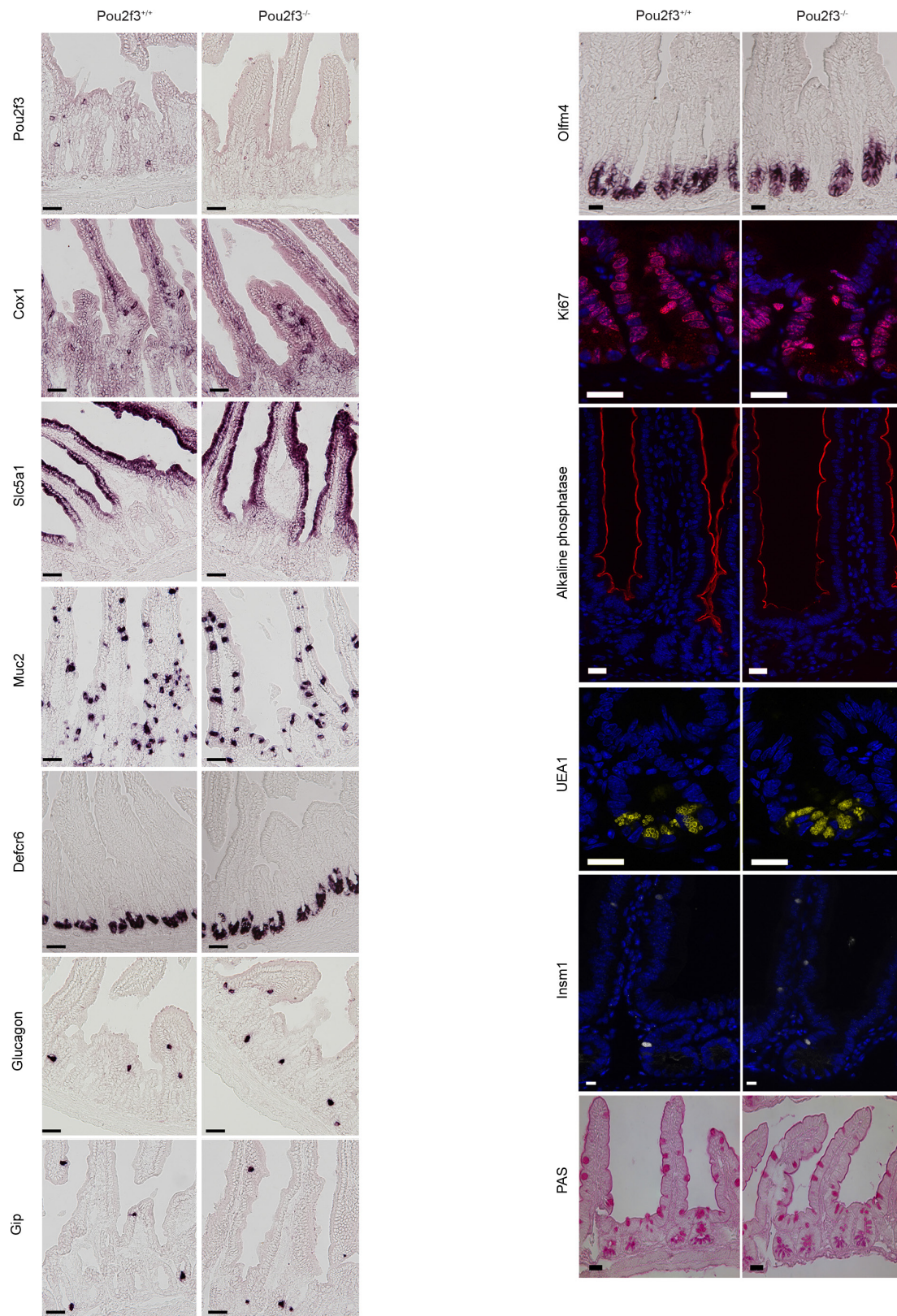
Extended Data Figure 1 | Epithelial differentiation parameters during *Nb* infection. **a**, Graph showing the distribution of Dclk1⁺ tuft cells in naive and infected mice 4, 5 and 7 days post infection. Cells were counted in the crypt and villus compartments of $n = 50$ crypt-villus units per mouse with 3 mice per condition. Means of villus/crypt ratio of tuft cell numbers are shown. **b**, Quantification of the goblet cell hyperplasia in naive and infected mice 4, 5 and 7 days post infection ($n = 50$ crypt-villus units per mouse; 3 mice per condition). **c**, Neo-differentiating tuft cells following *Nb* infection are indistinguishable from the tuft cells found in naive mouse intestinal epithelium as shown with Sox9 and Plc γ 2 stainings ($n = 3$ mice). **d**, Proliferation status of Pou2f3⁺ tuft cells in naive and

infected mice, shown with co-expression with the Ki67 proliferation marker. Arrows indicate Ki67⁺ cells located at various positions along the crypt axis. **e**, Increased proliferation of Pou2f3⁺ tuft cells during response to *Nb* infection ($n = 3$ naive and 3 infected mice). **f**, Dclk1⁺ tuft cells and Insm1⁺ enteroendocrine cells are distinct populations ($n = 3$ mice). **g**, Decrease of the Insm1⁺ enteroendocrine cell population during type 2 responses to *Nb* infection, concomitant to the expansion of the tuft cell lineage 7 days post infection ($n = 3$ naive and 3 infected mice). All the histograms show means \pm s.d. A two-tailed Student's *t*-test with Welch's correction was used, except for **g** where the 2 groups displayed comparable variances. All stainings were repeated 3 times.



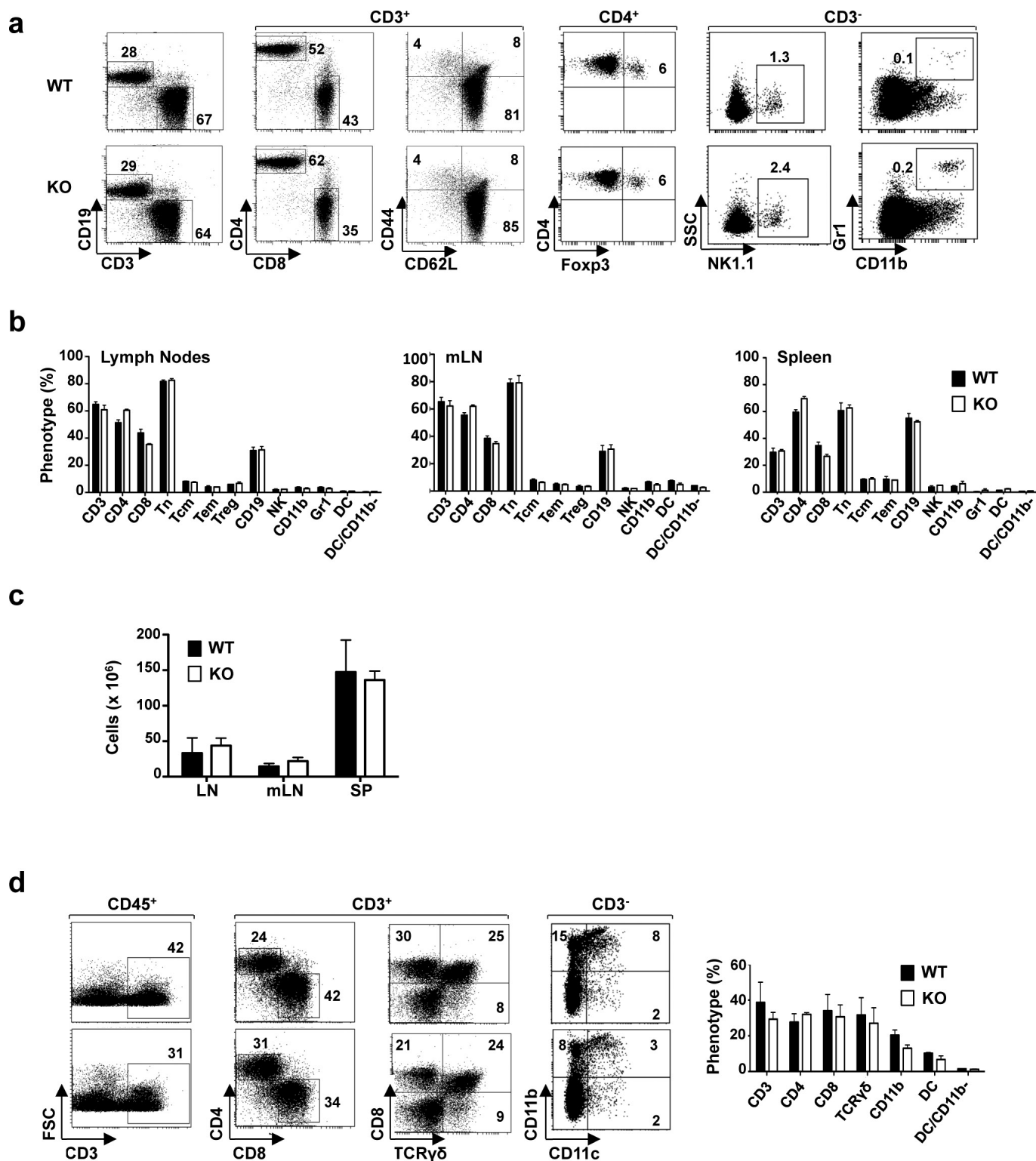
Extended Data Figure 2 | Expansion of the tuft cell lineage is a common adaptation of the intestinal epithelium following infection with helminth parasites. Tuft cell lineage expansion was assessed by Dclk1 immunohistochemistry in 2 different genetic backgrounds following infection with two different helminths, at the indicated time points. **a, b**, Naive and *Nb*-infected BALB/c mice. **c, d**, Naive and

H. polygyrus-infected C57BL/6 and BALB/c mice. **e, f**, Naive and *Nb*-infected C57BL/6 and Rag^{-/-} mice. **b, d, f**, $n = 50$ crypt-villus units per mouse; 3 mice per condition. Data are shown as means \pm s.d. and P values are indicated. A two-tailed Student's t -test with Welch's correction was used. Scale bars, 20 μ m. All experiments displayed in this figure were repeated 3 times.



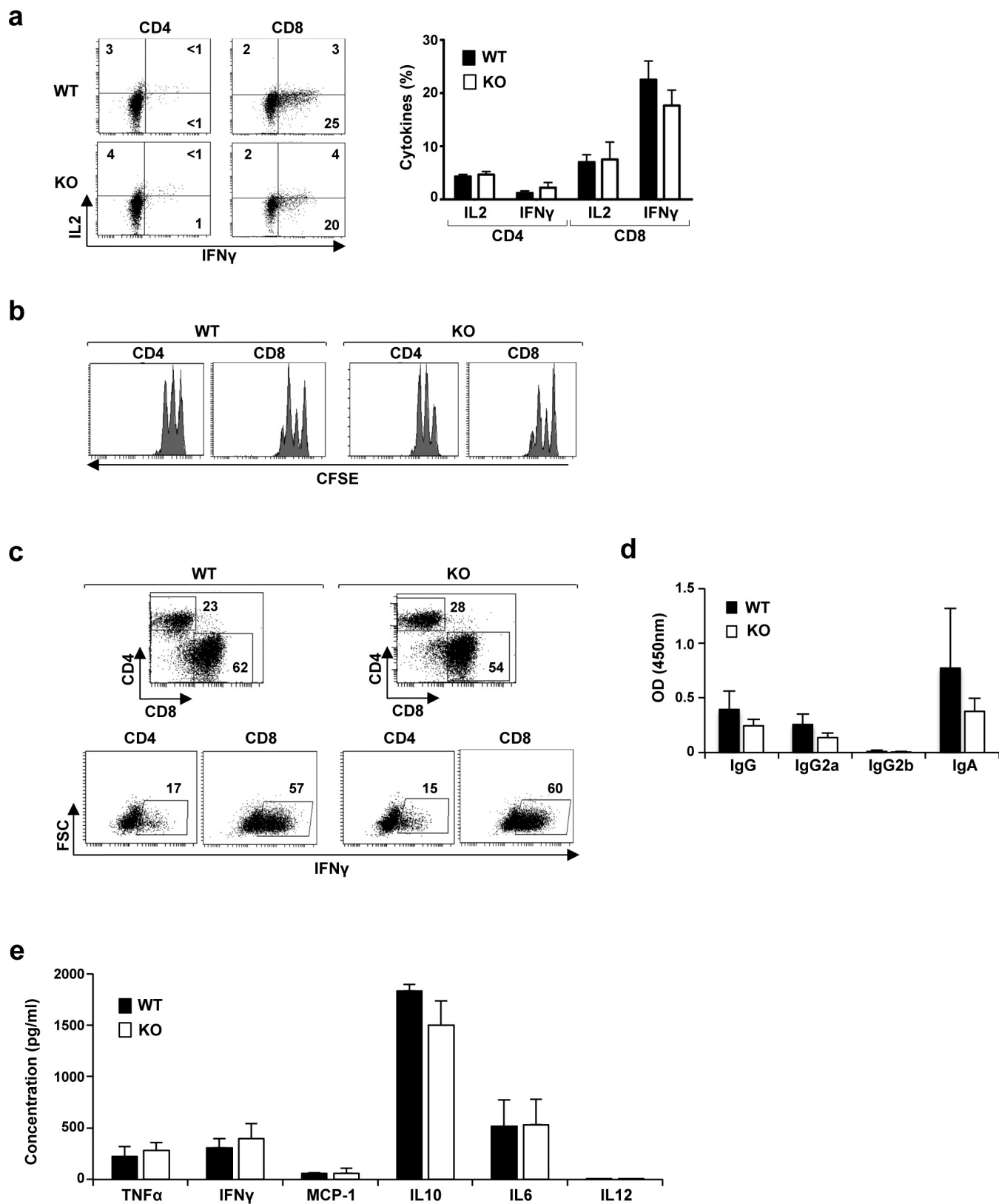
Extended Data Figure 3 | *Pou2f3* deficiency results in the specific absence of tuft cells in the intestinal epithelium. Characterization of the intestinal epithelium in *Pou2f3*-deficient mice as compared with wild-type littermate controls ($n = 3$ mice of each genotype). Left, *in situ* hybridization showing absence of tuft cells (*Pou2f3*, *Cox1*) in *Pou2f3*^{-/-} mice, whereas enterocytes (*Slc5a1*), goblet cells (*Muc2*), Paneth cells (*Defcr6*) and enteroendocrine cells (glucagon, *Gip*) are unaffected. Right, representative pictures of the *in situ* hybridization (*Olfm4*) and

immunohistochemistry experiments underlying the quantitative analysis provided in Fig. 2c showing that the stem cells (*Olfm4*), proliferative compartment (*Ki67*), and differentiated cell types: enterocytes (alkaline phosphatase), Paneth (*UEA1*), enteroendocrine (*Insm1*) and goblet (PAS staining) cells populations are unaffected in the *Pou2f3*^{-/-} mice. All panels show representative pictures of experiments replicated 3 times in 3 different mice. Scale bars, 20 μm.



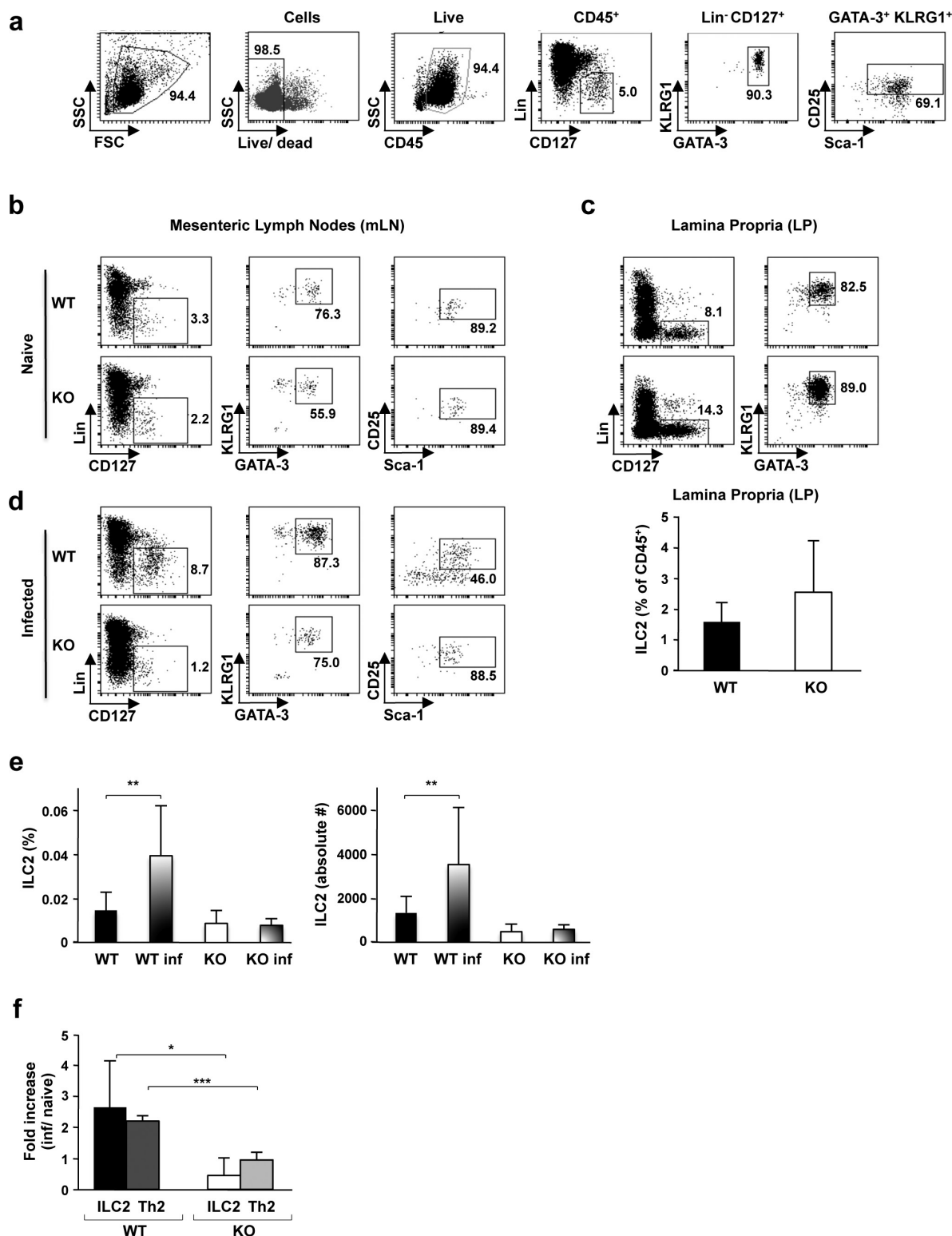
Extended Data Figure 4 | Immune cell homeostasis is not altered in *Pou2f3*^{-/-} tuft-cell-deficient mice. **a**, The repartition of immune cells in wild-type and *Pou2f3*^{-/-} mice was monitored by flow cytometry. The presence of T (CD3⁺), B (CD19⁺), CD4⁺, CD8⁺, naive (Tn; CD3⁺CD62L⁺CD44⁻), central memory (Tcm; CD3⁺CD62L⁺CD44⁺), effector memory (Tem; CD3⁺CD62L⁻CD44⁺), regulatory (Treg; CD4⁺Foxp3⁺), natural killer (NK; CD3⁻NK1.1⁺) and myeloid (CD11b⁺Gr1⁺) cells was assessed by staining with fluorochrome-tagged antibodies and representative dot plots are shown. The percentages of positively-stained cells are indicated. **b**, Quantification of the different immune cells in lymph nodes (LN), mesenteric lymph nodes (mLN) and spleens (SP) of wild-type and *Pou2f3*^{-/-} mice are presented. Data are means \pm s.d.

($n = 3$ mice per genotype). **c**, Total cells in LN, mLN and SP of wild-type and *Pou2f3*^{-/-} mice are presented as means \pm s.d. ($n = 3$ mice per genotype). **d**, Immune lineage cells in the lamina propria of wild-type and *Pou2f3*^{-/-} mice were monitored by flow cytometry after tissue dissociation. The percentage of T cells was assessed within the CD45⁺ haematopoietic gate, CD4, CD8 and gamma-delta T cells (CD8⁺TCR- $\gamma\delta$ ⁺) within the CD3⁺ gate and myeloid cells within the CD3⁻ gate, as indicated. Representative dot plots are presented (left). Quantification of immune cells within the lamina propria are shown as means \pm s.d. ($n = 3$ mice per group). No significant differences were detected for all cell types between wild-type and *Pou2f3*^{-/-} mice ($P > 0.05$). A two-tailed Student's *t*-test was used.



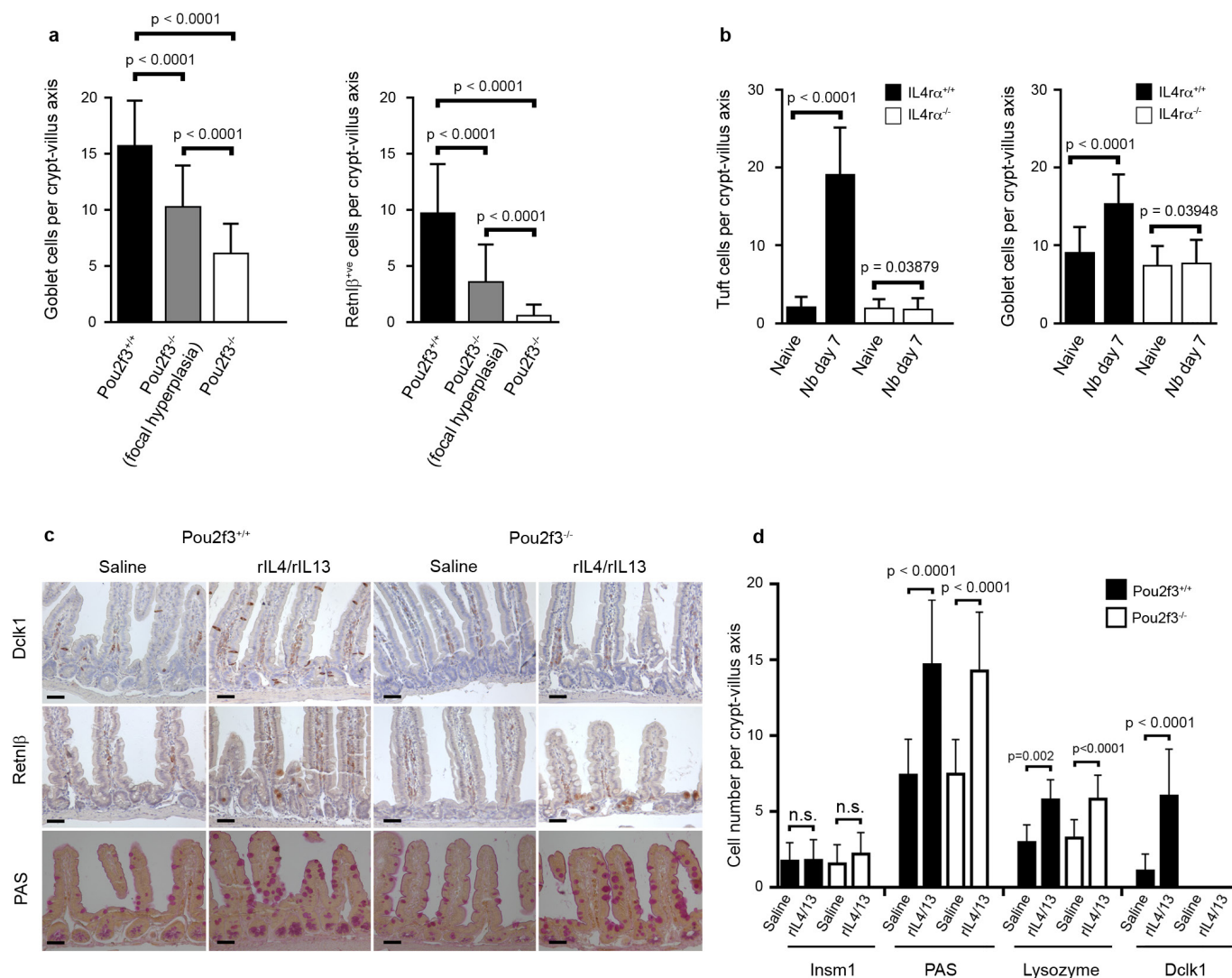
Extended Data Figure 5 | Equivalent immune responsiveness of wild-type and *Pou2f3*^{-/-} lymphocytes. **a**, The level of IL-2 and interferon gamma (IFN- γ) production by *Pou2f3*^{+/+} and *Pou2f3*^{-/-} CD4 and CD8 lymph node T cells was monitored directly after *ex vivo* isolation and representative histograms are presented (left). Quantification of cytokine secreting CD4 and CD8 T cells are presented as means \pm s.d. ($n = 3$ per group; $P > 0.05$). **b**, CFSE-loaded T cells were activated with immobilized anti-CD3/anti-CD28 antibodies for 2 days and proliferation was monitored as a function of fluorescence dilution. Representative

histograms for CD4 and CD8 T cells are shown. **c**, IFN- γ production in wild-type and *Pou2f3*^{-/-} lymphocytes was assessed at day 6 post CD3/CD28 stimulation and representative plots for CD4 and CD8 T cells are presented. **d**, Splenocytes from wild-type and *Pou2f3*^{-/-} mice were activated with LPS+IL-4 for 40 h and levels of secreted IgG, IgG2a, IgG2b and IgA were monitored by ELISA. Means \pm s.d. are presented. **e**, Splenocytes were activated as above and levels of TNF- α , IFN- γ , MCP-1, IL-10, IL-6, and IL-12 were monitored by cytometric bead array. Means \pm s.d. are presented. A two-tailed Student's *t*-test was used.



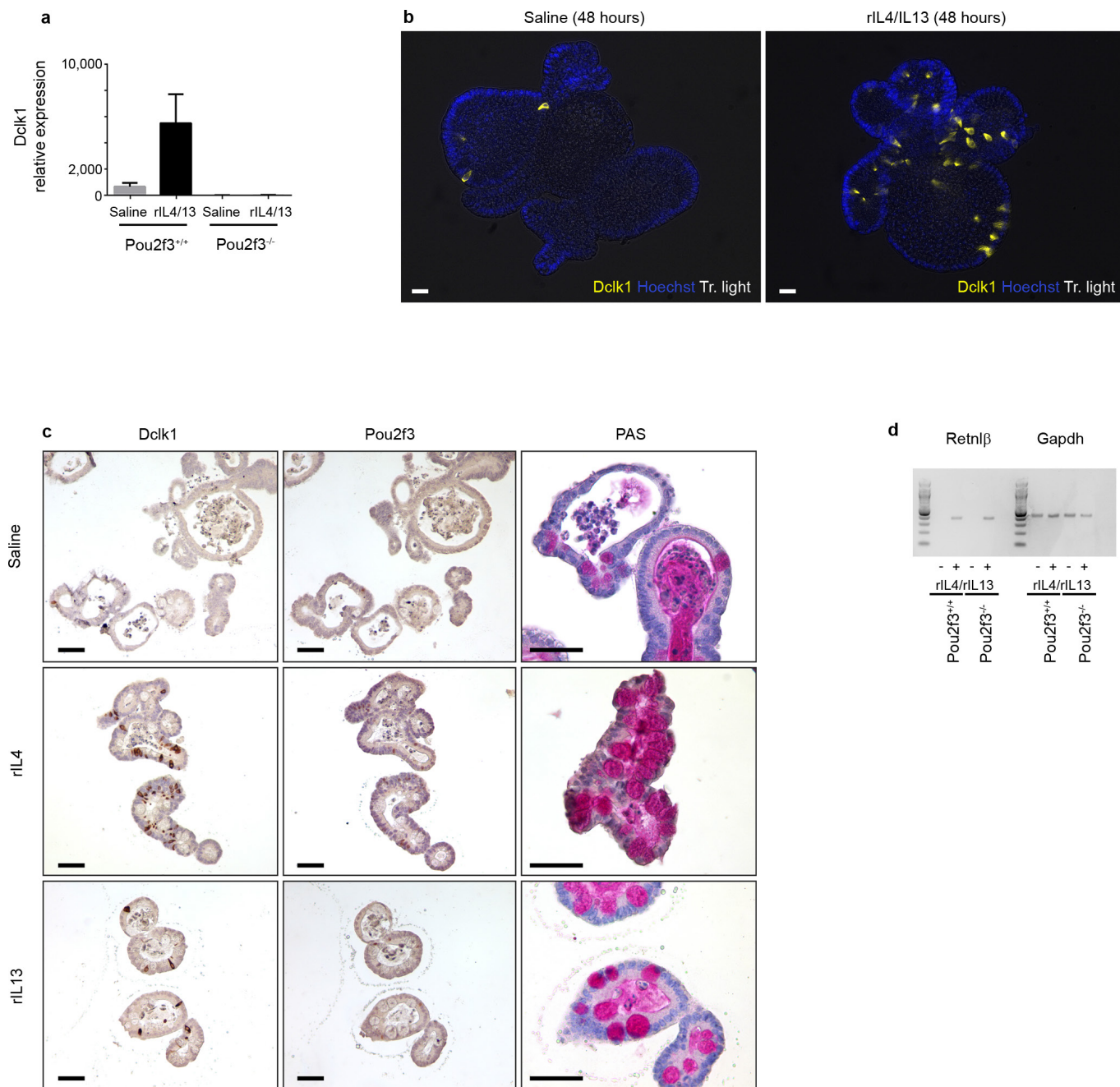
Extended Data Figure 6 | Defective induction of type 2 immunity in *Pou2f3*^{-/-} mice following helminth infection. **a**, Flow cytometry gating strategy for analysis of the innate ILC2 subset is shown. ILC2s were assessed within the CD45⁺ haematopoietic subset as lineage-CD127⁺ cells expressing KLRG1, GATA-3, Sca-1 and CD25 cell surface markers. Numbers represent the percentages of boxed cells. The staining strategy was validated using mLN cells from ZAP-70^{-/-} mice as this subset is present at relatively high levels in these immunodeficient mice. **b**, The presence of ILC2 cells in mLNs of naive *Pou2f3*^{+/+} (WT) and *Pou2f3*^{-/-} (KO) mice was assessed using the gating strategy shown above. Representative data from WT ($n=8$) and KO ($n=5$) mice are presented. **c**, Representative plots of ILC2 cells in lamina propria of naive

WT ($n=7$) and KO ($n=5$) mice are shown (top). Quantifications of ILC2 are presented as means \pm s.d. **d**, WT and KO mice were infected with *N. brasiliensis* and the presence of ILC2 in mLNs was assessed 5 days post infection. Representative plots are shown ($n=6$ mice per group). **e**, Quantification of ILC2 cells in mLN of naive versus infected WT and KO mice. The percentage of ILC2s within the live gate (left) and the absolute numbers of ILC2s (right) are presented. Data are means \pm s.d. ($n=5$ for WT, $n=8$ for KO, $n=6$ for both groups of infected mice). ** $P=0.01$. **f**, The fold-increase in ILC2 (lineage⁻CD127⁺KLRG1⁺GATA-3⁺) and Th2 (CD3⁺CD4⁺Gata-3⁺) cells in mLN was assessed as a function of infection ($n=6$ per group). The mean fold-increase \pm s.d. in WT and KO mice is presented. * $P=0.02$, *** $P=0.0005$. A two-tailed Student's *t*-test was used.



Extended Data Figure 7 | Signalling via IL-4R α is required and sufficient to induce goblet and tuft cell hyperplasia. a, Quantification of goblet (PAS and Retnl β staining) cells in *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice infected with *Nb* (day 7 post infection). In *Pou2f3*^{-/-} mice, crypt-villus axes from both focally responding regions and the rest of the tissue were counted. **b,** Quantification of tuft cells (Dclk1 staining) and goblet cell hyperplasia (PAS) in *Il4ra*^{+/+} and *Il4ra*^{-/-} mice. **c,** Histological analysis showing tuft (Dclk1 staining) and goblet (PAS and Retnl β staining) cells in *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice following treatment with a mixture

of rIL-4 and rIL-13 for 5 days. $n = 3$ mice per condition. All panels show representative experiments replicated 3 times. Scale bars, $20\mu\text{m}$. **d**, Quantitative analysis of the changes in the different cell types of the intestinal epithelium of *Pou2f3*^{+/+} and *Pou2f3*^{-/-} mice following treatment with a mixture of rIL-4 and rIL-13 during 5 days. For **a**, **b**, **d**, $n = 50$ crypt-villus axes counted in 3 mice per genotype or condition. Data are shown as means \pm s.d. and P values are indicated. A two-tailed Student's t -test with Welch's correction was used.

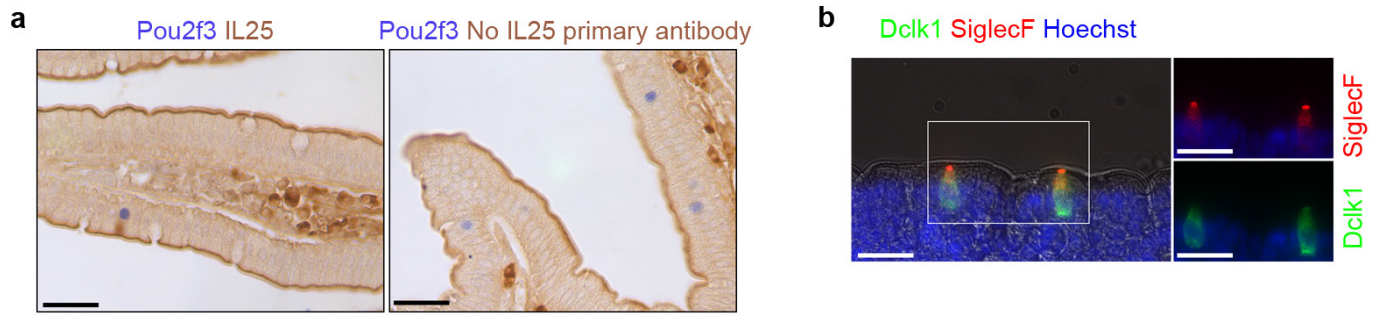


Extended Data Figure 8 | Signalling via IL-4R α is sufficient to induce goblet and tuft cell hyperplasia in mouse intestinal organoids.

a, Quantification of *Dclk1* expression analysis by qRT-PCR in *Pou2f3*^{+/+} and *Pou2f3*^{-/-} organoids following rIL-4/rIL-13 treatment for 48 h to assess the presence and amplification of tuft cells. Means \pm s.d., relative to *Gapdh* and *Hprt*, are presented. **b**, Expansion of the tuft cell lineage in wild-type organoids following rIL-4/rIL-13 administration (48 h)

was monitored by Dclk1 staining. **c**, Expansion of the tuft cell lineage in wild-type organoids following IL-4 or IL-13 administration (48 h) was monitored by Dclk1, Pou2f3 and PAS stainings. Scale bars, 20 μ m.

d, *Retnlβ* expression in *Pou2f3*^{+/+} and *Pou2f3*^{-/-} organoids was monitored as a function of rIL-4/rIL-13 treatment (48 h) by RT-PCR and data relative to *Gapdh* are presented. All panels show representative experiments from 3 independent organoid cultures, replicated 3 times.



Extended Data Figure 9 | Validation of the Siglec-F and IL-25 stainings.
a, Control experiment for specificity of the IL-25 immunohistochemistry, in presence (left) or absence (right) of IL-25 primary antibody.

b, Immunohistochemistry showing specificity of Siglec-F as a marker for intestinal epithelial tuft cells. All panels show representative experiments from 3 independent mice, replicated 3 times.

Extended Data Table 1 | List of the oligonucleotide primer sequences

Table 1: Primer sets used for qPCR analyses (5' 3'):

Dclk1	CAGCCTGGACGAGCTGGTGG	TGACCAGTTGGGGTTTCACAT
Gapdh	GGAGCGAGACCCCACTAACA	ACATACTCAGCACCGGCCTC
Hprt	GCAGTACAGCCCCAAAATGG	GGTCCTTTTCACCAGCAAGCT

Table 2: Primer sets used for PCR analyses (5' 3'):

IL13	AGCTCCCTGGTTCTCTCACT	CTCATTAGAAGGGGCCGTGG
IL25	TCTTGGCAATGATCGTGGGA	TGTGGTAAAGTGGGACGGAG
Retn1b	CCAGAAGACCATTTCTTGAGCT	CCCACGATCCACAGCCATAG
Gapdh	CAAGAAGGTGGTGAAGCAGG	TCAAGAGAGTAGGGAGGGCT
Pou2f3	AGAGAATCAACTGCCCCGTG	GGAAGGCACGACTCTCTTCC

Crystal structure of a DNA catalyst

Almudena Ponce-Salvatierra^{1,2}, Katarzyna Wawrzyniak-Turek^{1,3}, Ulrich Steuerwald², Claudia Höbartner^{1,3} & Vladimir Pena²

Catalysis in biology is restricted to RNA (ribozymes) and protein enzymes, but synthetic biomolecular catalysts can also be made of DNA (deoxyribozymes)¹ or synthetic genetic polymers². *In vitro* selection from synthetic random DNA libraries identified DNA catalysts for various chemical reactions beyond RNA backbone cleavage³. DNA-catalysed reactions include RNA and DNA ligation in various topologies^{4,5}, hydrolytic cleavage^{6,7} and photorepair of DNA⁸, as well as reactions of peptides^{9,10} and small molecules^{11,12}. In spite of comprehensive biochemical studies of DNA catalysts for two decades, fundamental mechanistic understanding of their function is lacking in the absence of three-dimensional models at atomic resolution. Early attempts to solve the crystal structure of an RNA-cleaving deoxyribozyme resulted in a catalytically irrelevant nucleic acid fold¹³. Here we report the crystal structure of the RNA-ligating deoxyribozyme 9DB1 (ref. 14) at 2.8 Å resolution. The structure captures the ligation reaction in the post-catalytic state, revealing a compact folding unit stabilized by numerous tertiary interactions, and an unanticipated organization of the catalytic centre. Structure-guided mutagenesis provided insights into the basis for regioselectivity of the ligation reaction and allowed remarkable manipulation of substrate recognition and reaction rate. Moreover, the structure highlights how the specific properties of deoxyribose are reflected in the backbone conformation of the DNA catalyst, in support of its intricate three-dimensional organization. The structural principles underlying the catalytic ability of DNA elucidate differences and similarities in DNA versus RNA catalysts, which is relevant for comprehending the privileged position of folded RNA in the prebiotic world and in current organisms.

The deoxyribozyme 9DB1 catalyses the regioselective formation of a native phosphodiester bond between the 3'-hydroxyl and the 5'-triphosphate termini of two RNA strands, using divalent metal ions (Mg^{2+} or Mn^{2+}) as cofactors^{14,15}. Ribozymes that catalyse an analogous ligation reaction are found in nature¹⁶, and have also been generated by *in vitro* selection from random RNA libraries^{17,18} or a natural ribozyme scaffold¹⁹. The 9DB1 DNA enzyme is composed of a central catalytic domain flanked by two arms that hybridize to the RNA substrates by canonical Watson–Crick base pairing (Fig. 1a, b), leaving only the two nucleotides embracing the ligation junction unpaired. The *in vitro* selection strategy used to obtain 9DB1 imposed selection pressure for base pairing outside the ligation junction and enforced 3',5' selectivity of the reaction¹⁴.

We crystallized a 44-nucleotide strand that contains the minimally required catalytic core of 31 nucleotides²⁰ in complex with a 15-nucleotide RNA strand. The structure was solved by single anomalous dispersion using data collected from hexamine cobalt (III) chloride. The model built in the experimental map was refined against a 2.8 Å resolution data set collected from a native crystal (Fig. 1c). The crystal contains one molecule in the asymmetric unit and the final model comprises all nucleotides present in the complex (Extended Data Table 1).

The structure exhibits the shape of a λ letter, consisting of the two RNA–DNA hybrid duplexes, P1 and P4, oriented at approximately 120° with respect to each other, and both tightly anchored to

the catalytic domain (Fig. 1d, e). The RNA nucleotides A–1 and G1 encompass the ligation junction and form extensive tertiary contacts with the DNA catalyst (Fig. 1d, e). The 3'-overhanging DNA and RNA nucleotides form a semi-continuous duplex in the crystal lattice

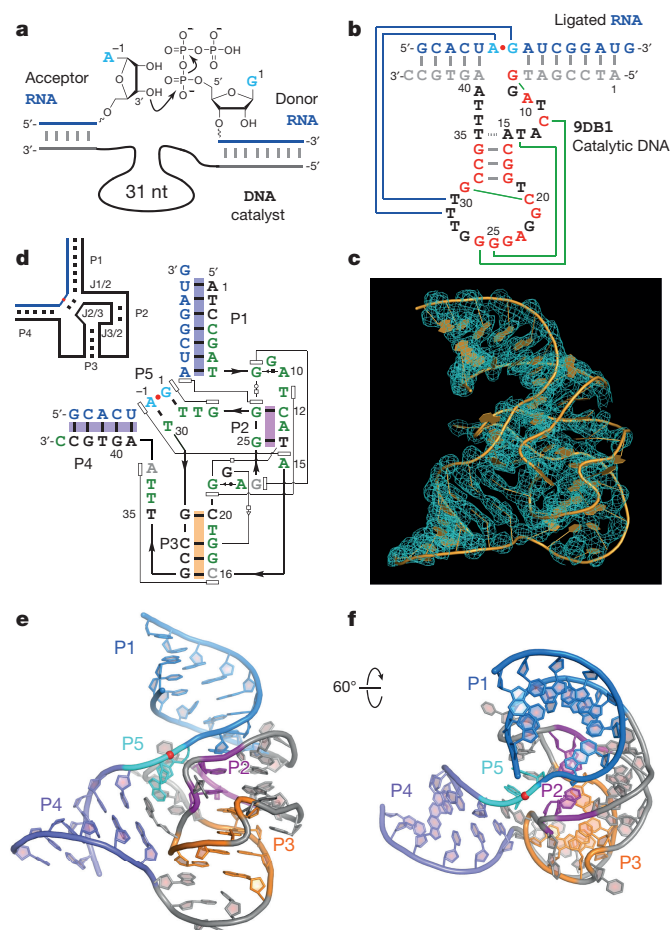


Figure 1 | Global architecture of the DNA catalyst. **a**, DNA-catalysed ligation between 3'-OH and 5'-phosphate. nt, nucleotides. **b**, Proposed secondary structure of the minimal 9DB1 deoxyribozyme²⁰. RNA nucleotides are in blue, ligation junction in light blue, DNA-binding arms in grey, and core nucleotides in black and red. Nucleotides in red do not tolerate any mutation. Blue and green lines indicate double pseudoknot interactions. **c**, Solvent-flattened electron density map contoured at 1.0σ level with orange trace of backbone and nucleobases. **d**, Secondary and tertiary structure with base pairs denoted in Leontis–Westhof presentation. DNA nucleotides in black have N-type ribose pucker, those in green have S-type conformations; grey nucleotides lie outside typical N or S conformations. P1–P5 are colour coded with respect to the structure shown in **e**. Inset: schematic overview of the double pseudoknot. **e**, Cartoon representation of the deoxyribozyme in complex with the ligated RNA product. **f**, A 60° rotation of the image shown in **e**.

¹Max Planck Research Group Nucleic Acid Chemistry, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. ²Research Group Macromolecular Crystallography, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. ³Institute for Organic and Biomolecular Chemistry, Georg-August-University Göttingen, Tammannstr. 2, 37077 Göttingen, Germany.

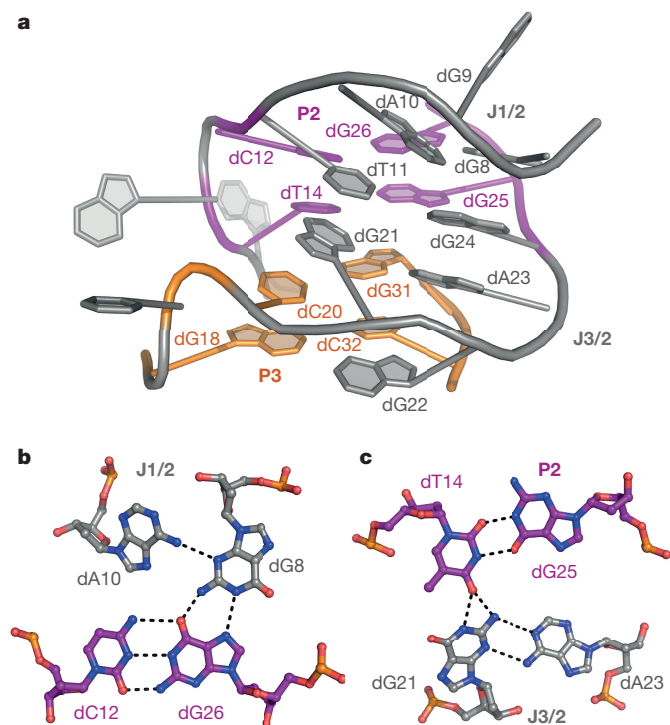


Figure 2 | Tertiary contacts within the catalytic domain. **a**, Multiple stacking and base pairing contacts in the 9DB1 core. **b**, **c**, Close-up view of multiple P2 and P3 contacts. Black dashes indicate hydrogen bonds.

(Extended Data Fig. 1). Within the catalytic domain, two stacks of two and four base pairs are formed, called P2 and P3, respectively. The paired regions P1, P2 and P3 are positioned on top of one another in a non-coaxial manner (Fig. 1f). In addition, the DNA residues dT29–dT30 base pair with the RNA nucleotides A–1–G1 (denoted P5), which makes the entire DNA–RNA complex fold as a double pseudoknot (Fig. 1d).

The overall fold of the catalytic domain is stabilized by tertiary interactions that connect the regions labelled J1/2, J3/2, P2 and P3 in four successive layers of multiplets (Fig. 2a). Hydrogen bonding forms the long-range Watson–Crick base pair dC12:dG26 and the non-canonical Watson–Crick base pair dT14:dG25 in P2 (Fig. 2b, c), as well as alternative base pairs involving the sugar edges of dG8 and dG21, the Watson–Crick edges of dG8, dG21 and dA23, and the Hoogsteen edges of dA10, dG22 and dG26. Moreover, the compact DNA architecture is stabilized by stacking interactions of nucleotides in J1/2 and J3/2, which place dT11 and dG24 sandwiched between the non-canonical base pairs dG8:dA10 and dG21:dA23. These interactions are consistent with previous atomic mutagenesis data, which showed that the Watson–Crick edges of dG8, dT14, dG21 and dG25 are essential for the functional integrity of 9DB1 (ref. 21). Moreover, dG26 did not tolerate changes on the Watson–Crick or Hoogsteen edges, consistent with its interactions in a base triple with the dC12 and dG8 nucleotides (Fig. 2b).

The reactive nucleotides A–1 and G1 are positioned in the active site by a scaffold formed by four nucleotides from distinct locations of the primary sequence, corroborated by the large angular opening between the duplexes P1 and P4 (Fig. 3a, b). Thus, the junction J2/3 protrudes from the catalytic domain and extensively contacts A–1 and G1 by base pairing with dT30 and dT29, respectively, as well as by the stacking of dG27 on G1 (Fig. 3a, b). On the opposite side, a 130° kink in the backbone between dT14 and dC16 causes dA15 to form a stacking interaction with A–1 (Fig. 3b). In this way, A–1 and G1 are stacked on one another, sandwiched by dA15 and dG27 and engaged in hydrogen bonding with dT29 and dT30, resulting in a microenvironment reminiscent of a duplex.

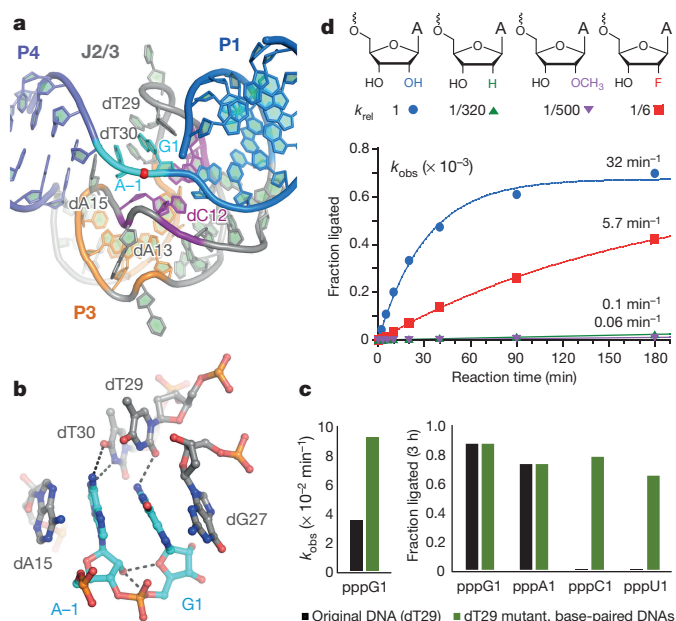


Figure 3 | Positioning of reactive nucleotides in the catalytic domain.

a, Structure at the ligation junction. **b**, Base-pairing and stacking interactions of the ligated nucleotides G1 and A–1. Black dashes indicate hydrogen bonds. **c**, Biochemical interrogation of the base pair of nucleotides 1:29. Ligation rate or yield with original 9DB1 (dT29, black), and with mutant 9DB1 DNAs (green) complementary to the RNA substrates, that is, dT29dC for pppC1, dT29dA for pppU1 RNA. **d**, Activity assays of 9DB1 with substrate analogues modified at the 2'-position of A–1. **c**, **d**, Single-turnover kinetic assays were performed at pH 7.5 with 20 mM Mn^{2+} at 37 °C. All experiments were repeated two or three times. k_{obs} , observed rate constant, k_{rel} , relative rate constant with respect to A–1 substrate.

The identification of the previously unknown base pairs (P5) at the ligation junction was intriguing, as dT29 and dT30 were shown to tolerate mutations²⁰, and RNA substrates with uridine or guanosine at position –1 were also efficiently ligated^{14,15}. In contrast, only 5'-purine nucleotides were allowed at position 1 in the donor RNA (that is, 5'-pppG1 or 5'-pppA1)¹⁴. Therefore, the crystal structure of 9DB1 revealed an effective way of manipulating the catalytic activity by compensatory mutagenesis of the base pair at the ligation junction. Indeed, RNA substrates that were inert to the original deoxyribozyme, such as 5'-pppC1 and 5'-pppU1 RNA, are now readily ligated with mutated 9DB1 enzymes bearing G and A, respectively, in position 29 (Fig. 3c). On the basis of the crystal structure, 9DB1 deoxyribozymes can now be designed for ligation of any RNA substrates, irrespective of their sequence.

The regioselectivity of the ligation reaction is determined by the relative orientation of the reacting nucleotides. For ribozymes^{22,23} as well as deoxyribozymes²⁴, continuous Watson–Crick base pairing across the ligation junction favours nucleophilic attack by the 3'-OH group onto the 5'-triphosphate of G1. The regioselectivity of 9DB1 is dictated by tertiary contacts and specific interactions at the ligation junction, which differ markedly from a continuous double-helical arrangement. Thus, the orientation of the donor nucleotide G1 is assisted by a hydrogen bond of its 2'-OH to the minor groove of dC12:dG26 in P2 (Fig. 4a). The arrangement of the binding arms induces a shorter distance between the phosphate groups of A–1 and G1 (5.3 Å as compared to 5.9 Å in a regular A-form duplex). The hydrogen bond between the 2'-OH group of A–1 and the O4' (or O5') of G1 may sequester the 2'-OH and thereby assist the regioselective formation of 3'–5' ligated products (Fig. 3b). Consistently, removal or blocking of the 2'-OH group by introducing a 2'-deoxy or 2'-OCH₃ nucleotide at position A–1 prevented efficient ligation (Fig. 3d). Instead, replacement of the 2'-OH with 2'-NH₂ still allowed RNA ligation, although with a 30-fold

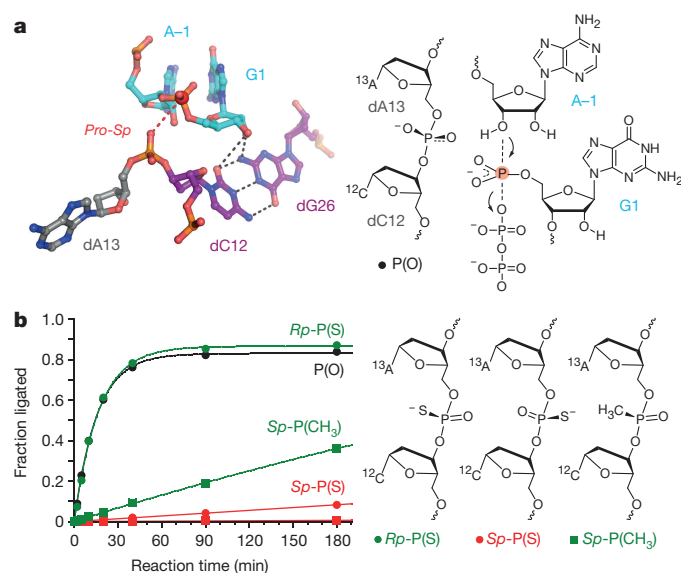


Figure 4 | Active site of 9DB1. **a**, The ligation junction and the putatively catalytic dA13 phosphate are depicted based on the crystal structure (left) and as schematic drawing of the ligation reaction (right). Red dashes indicate the proximity between *pro*-Sp oxygen and phosphorous atom of newly formed phosphodiester bond (red ball). **b**, Biochemical interrogation of the phosphate participation by phosphorothioate and methylphosphonate interference (pH 7.5, 20 mM Mn²⁺). Diastereomers are schematically depicted. All measurements were repeated at least twice.

slower rate (Extended Data Fig. 2). Interestingly, 2'-fluoro-modified RNA was ligated only 6–7-fold slower (Fig. 3d and Extended Data Fig. 2), suggesting that the electronegativity of fluorine and the resulting 3'-endo conformation of the ribose compensates for the loss of the 2'-OH group and orients the 3'-OH for nucleophilic attack onto the 5'-phosphate.

Catalysis of phosphoryl-transfer reactions by proteins or ribozymes often involves activation of nucleophilic groups and electrostatic stabilization of the transition state by divalent metal ions or specific chemical groups^{25,26}. The active site of an RNA enzyme that catalyses the same reaction as the DNA enzyme 9DB1, is composed of two phosphate groups coordinating a catalytic metal ion, plus an amino group from a nucleobase and a 2'-OH group that exerts its action through intermediacy of a water molecule²⁷. There are no equivalent interactions in the core of 9DB1, and we do not observe electron density for a catalytic metal ion. However, a striking feature is the phosphate group of nucleotide dA13, which lies within 3.1 Å distance from the ligation junction (Fig. 4a). To test the role of the dA13 phosphate in catalysis, we replaced either of the non-bridging oxygen atoms with a sulfur atom, and assessed the enzymatic activity of the resulting phosphorothioate-modified DNAs. In the presence of manganese ions, the ligation rate of the deoxyribozyme containing an *Sp* phosphorothioate was reduced 100-fold, whereas the *Rp* stereoisomer retained its full activity (Fig. 4b). A similar phosphorothioate interference effect was observed when magnesium was used as metal ion cofactor, and the catalytic activity could also not be rescued by the addition of cadmium ions. The absence of thiophilic metal rescue cannot prove whether metal ions are involved or not in catalysis. Upon substitution of the non-bridging phosphate oxygens with a methyl group, only one of the two resulting uncharged methylphosphonate diastereomers retained catalytic activity (Fig. 4b). Although the role of metal ions remains unclear, these data are consistent with the critical role of the non-bridging *pro*-Sp phosphate oxygen of dA13 in activating the 3'-OH group to promote RNA ligation.

This work provides the experimental evidence that DNA possesses the ability to fold into compact three-dimensional units, stabilized by intrinsic tertiary interactions in the absence of 2'-OH groups or

artificial hydrophobic substituents²⁸. In contrast with ribozymes, the sugar-phosphate backbone of the catalytic domain of 9DB1 exhibits much larger conformational diversity, as reflected by the broad distribution of pseudorotation phase angles found in the DNA nucleotides (Extended Data Fig. 3). Thus, of the 31 nucleotides in the catalytic domain, 8 and 20 adopt north (N)- and south (S)-type ribose puckers, respectively (Fig. 1d). The remaining three nucleotides lie outside typical N or S conformations. The less restrictive puckering allows DNA to explore a wide range of conformations, compensating for the lack of the 2'-hydroxyl group, which has an important structural role in ribozymes.

Comparison of 9DB1 with two RNA ligase ribozymes illustrates how the fold of DNA and RNA enzymes underlies the recognition and positioning of the substrates in the catalytic centre (Extended Data Fig. 4). In the class I ligase and L1 ligase ribozymes, the reactive nucleotides are positioned in extended helices (Extended Data Fig. 5)^{29,30}. In contrast, the four contributing bases in the 9DB1 ligase are provided *in trans* from different primary locations of the catalytic domain. Notably, all four DNA nucleotides adopt S-type ribose puckers, highlighting how the specific conformational propensity of deoxyribose moieties supports recognition and positioning of the substrate in the active centre.

The crystal structure of the 9DB1 deoxyribozyme demonstrates that DNA possesses the intrinsic ability to adopt complex tertiary folds that support catalysis, and unveils for the first time the active site of a DNA enzyme in the post-catalytic state. Together with mutagenic analyses, the structure elucidates the basis of substrate recognition and provides first insights for rationalization of the regiospecific bond formation. The ability of DNA to form complex three-dimensional architectures that support catalysis raises questions about the positioning of this biopolymer in the prebiotic evolution of life and may shed light on metabolic events in current organisms, in which single-stranded DNA may adopt functionally important folds.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 July; accepted 20 November 2015.

Published online 6 January 2016.

- Breaker, R. R. & Joyce, G. F. A DNA enzyme that cleaves RNA. *Chem. Biol.* **1**, 223–229 (1994).
- Taylor, A. I. *et al.* Catalysts from synthetic genetic polymers. *Nature* **518**, 427–430 (2015).
- Schlosser, K. & Li, Y. Biologically inspired synthetic enzymes made from DNA. *Chem. Biol.* **16**, 311–322 (2009).
- Coppins, R. L. & Silverman, S. K. A DNA enzyme that mimics the first step of RNA splicing. *Nature Struct. Mol. Biol.* **11**, 270–274 (2004).
- Sreedhara, A., Li, Y. & Breaker, R. R. Ligating DNA with DNA. *J. Am. Chem. Soc.* **126**, 3454–3460 (2004).
- Chandra, M., Sachdeva, A. & Silverman, S. K. DNA-catalyzed sequence-specific hydrolysis of DNA. *Nature Chem. Biol.* **5**, 718–720 (2009).
- Gu, H., Furukawa, K., Weinberg, Z., Berenson, D. F. & Breaker, R. R. Small, highly active DNAs that hydrolyze DNA. *J. Am. Chem. Soc.* **135**, 9121–9129 (2013).
- Chinnappen, D. J. & Sen, D. A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl Acad. Sci. USA* **101**, 65–69 (2004).
- Walsh, S. M., Sachdeva, A. & Silverman, S. K. DNA catalysts with tyrosine kinase activity. *J. Am. Chem. Soc.* **135**, 14928–14931 (2013).
- Chandrasekar, J. & Silverman, S. K. Catalytic DNA with phosphatase activity. *Proc. Natl Acad. Sci. USA* **110**, 5315–5320 (2013).
- Li, Y. & Sen, D. A catalytic DNA for porphyrin metallation. *Nature Struct. Biol.* **3**, 743–747 (1996).
- Chandra, M. & Silverman, S. K. DNA and RNA can be equally efficient catalysts for carbon–carbon bond formation. *J. Am. Chem. Soc.* **130**, 2936–2937 (2008).
- Nowakowski, J., Shim, P. J., Prasad, G. S., Stout, C. D. & Joyce, G. F. Crystal structure of an 82-nucleotide RNA–DNA complex formed by the 10–23 DNA enzyme. *Nature Struct. Biol.* **6**, 151–156 (1999).
- Purtha, W. E., Coppins, R. L., Smalley, M. K. & Silverman, S. K. General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages. *J. Am. Chem. Soc.* **127**, 13124–13125 (2005).
- Büttner, L., Seikowski, J., Wawrzyniak, K., Ochmann, A. & Höbartner, C. Synthesis of spin-labeled riboswitch RNAs using convertible nucleosides and DNA-catalyzed RNA ligation. *Bioorg. Med. Chem.* **21**, 6171–6180 (2013).
- Vicens, Q. & Cech, T. R. A natural ribozyme with 3'/5' RNA ligase activity. *Nature Chem. Biol.* **5**, 97–99 (2009).

17. Bartel, D. P. & Szostak, J. W. Isolation of new ribozymes from a large pool of random sequences [see comment]. *Science* **261**, 1411–1418 (1993).
18. Ekland, E. H., Szostak, J. W. & Bartel, D. P. Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science* **269**, 364–370 (1995).
19. Jaeger, L., Wright, M. C. & Joyce, G. F. A complex ligase ribozyme evolved *in vitro* from a group I ribozyme domain. *Proc. Natl Acad. Sci. USA* **96**, 14712–14717 (1999).
20. Wachowius, F., Javadi-Zarnaghi, F. & Höbartner, C. Combinatorial mutation interference analysis reveals functional nucleotides required for DNA catalysis. *Angew. Chem. Int. Ed. Engl.* **49**, 8504–8508 (2010).
21. Wachowius, F. & Höbartner, C. Probing essential nucleobase functional groups in aptamers and deoxyribozymes by nucleotide analogue interference mapping of DNA. *J. Am. Chem. Soc.* **133**, 14888–14891 (2011).
22. Ikawa, Y., Tsuda, K., Matsumura, S. & Inoue, T. *De novo* synthesis and development of an RNA enzyme. *Proc. Natl Acad. Sci. USA* **101**, 13750–13755 (2004).
23. Pitt, J. N. & Ferré-D'Amaré, A. R. Structure-guided engineering of the regioselectivity of RNA ligase ribozymes. *J. Am. Chem. Soc.* **131**, 3532–3540 (2009).
24. Coppins, R. L. & Silverman, S. K. Rational modification of a selection strategy leads to deoxyribozymes that create native 3'–5' RNA linkages. *J. Am. Chem. Soc.* **126**, 16426–16432 (2004).
25. Doudna, J. A. & Cech, T. R. The chemical repertoire of natural ribozymes. *Nature* **418**, 222–228 (2002).
26. Fedor, M. J. & Williamson, J. R. The catalytic diversity of RNAs. *Nature Rev. Mol. Cell Biol.* **6**, 399–412 (2005).
27. Shechner, D. M. & Bartel, D. P. The structural basis of RNA-catalyzed RNA polymerization. *Nature Struct. Mol. Biol.* **18**, 1036–1042 (2011).
28. Davies, D. R. *et al.* Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proc. Natl Acad. Sci. USA* **109**, 19971–19976 (2012).
29. Shechner, D. M. *et al.* Crystal structure of the catalytic core of an RNA-polymerase ribozyme. *Science* **326**, 1271–1275 (2009).
30. Robertson, M. P. & Scott, W. G. The structural basis of ribozyme-catalyzed RNA assembly. *Science* **315**, 1549–1553 (2007).

Acknowledgements This work was supported by the Max Planck Society. We thank J. Ludwig for a gift of 5'-triphosphorylated RNAs, J. Seikowski for assistance with RNA and DNA synthesis, F. Wachowius and B. Samanta for initial crystallization samples, P. Afonine and members of the Pena and Höbartner laboratories for discussions, and the beamline staff at the Swiss Light Source, Villigen, Switzerland for assistance with data collection.

Author Contributions Crystallographic work was performed by A.P.-S. under the supervision of V.P., biochemical experiments were performed by A.P.-S., K.W.-T. and C.H. U.S. obtained initial crystals. A.P.-S., V.P. and C.H. designed the experiments and all authors discussed the results and contributed to the manuscript.

Author Information Structural models and structure factors have been deposited in the Protein Data Bank under accession numbers 5CKI (cobalt) and 5CKK (native). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.P. (vpenna@gwdg.de) or C.H. (choebar1@gwdg.de).

METHODS

Oligonucleotide synthesis and complex formation. RNA and DNA oligonucleotides were prepared by solid-phase synthesis using phosphoramidite chemistry on polystyrene or controlled-pore glass solid supports. For RNA synthesis 2'-O-TOM-protected RNA phosphoramidites were used³¹. 3'-Modified RNA was prepared on commercially available modified (2'-OCH₃, 2'-NH₂, 2'-F-uridine) solid supports or universal support (2'-F-adenosine). All oligonucleotides were deprotected with ammonia or methyl amine, followed by tetrabutylammonium fluoride (for RNA only), desalted and purified by denaturing polyacrylamide gel electrophoresis. The phosphorothioate modification was introduced during solid-phase synthesis using phenylacetyl disulfide³². Methylphosphonate modifications were introduced using commercially available N⁴-acetyl-dC methylphosphonamidite. Phosphorothioate and methylphosphonate diastereomers were separated by RP-C18 high-performance liquid chromatography (HPLC; NH₄OAc/CH₃CN gradient) and assigned according to elution order as reported^{33,34}. Full-length deoxyribozymes were generated by splinted ligation of two fragments. 5'-Triphosphorylated donor RNA with 5'-purine nucleotides were prepared by *in vitro* transcription with T7 RNA polymerase. Synthetic triphosphorylated RNAs were prepared by solid-phase synthesis following published procedures^{35,36} or provided by J. Ludwig. Crystallization complexes were formed by mixing RNA and DNA strands in a 1:1 molar ratio, heated to 95 °C, and slowly cooled down to room temperature before MgCl₂ was added. The final concentration of the complex was 0.8 mM in 10 mM HEPES pH 8.0, 150 mM NaCl, 2 mM KCl and 20 mM MgCl₂.

Crystallization and diffraction experiments. Crystals were grown at 4 °C from solutions containing 35% 2-methyl-2,5-pentanediol (MPD), 0.5 mM spermine, 20 mM MgCl₂, 100 mM NaCl and 100 mM sodium cacodylate (pH 6.5) using the sitting-drop vapour-diffusion method. Sitting drops with volumes of 2 µl produced crystals after 5–10 days, reaching full size after 30 days. The derivatized crystals were prepared by soaking of the crystals in a solution containing the mother liquor and hexammine cobalt (III) chloride for 1 min.

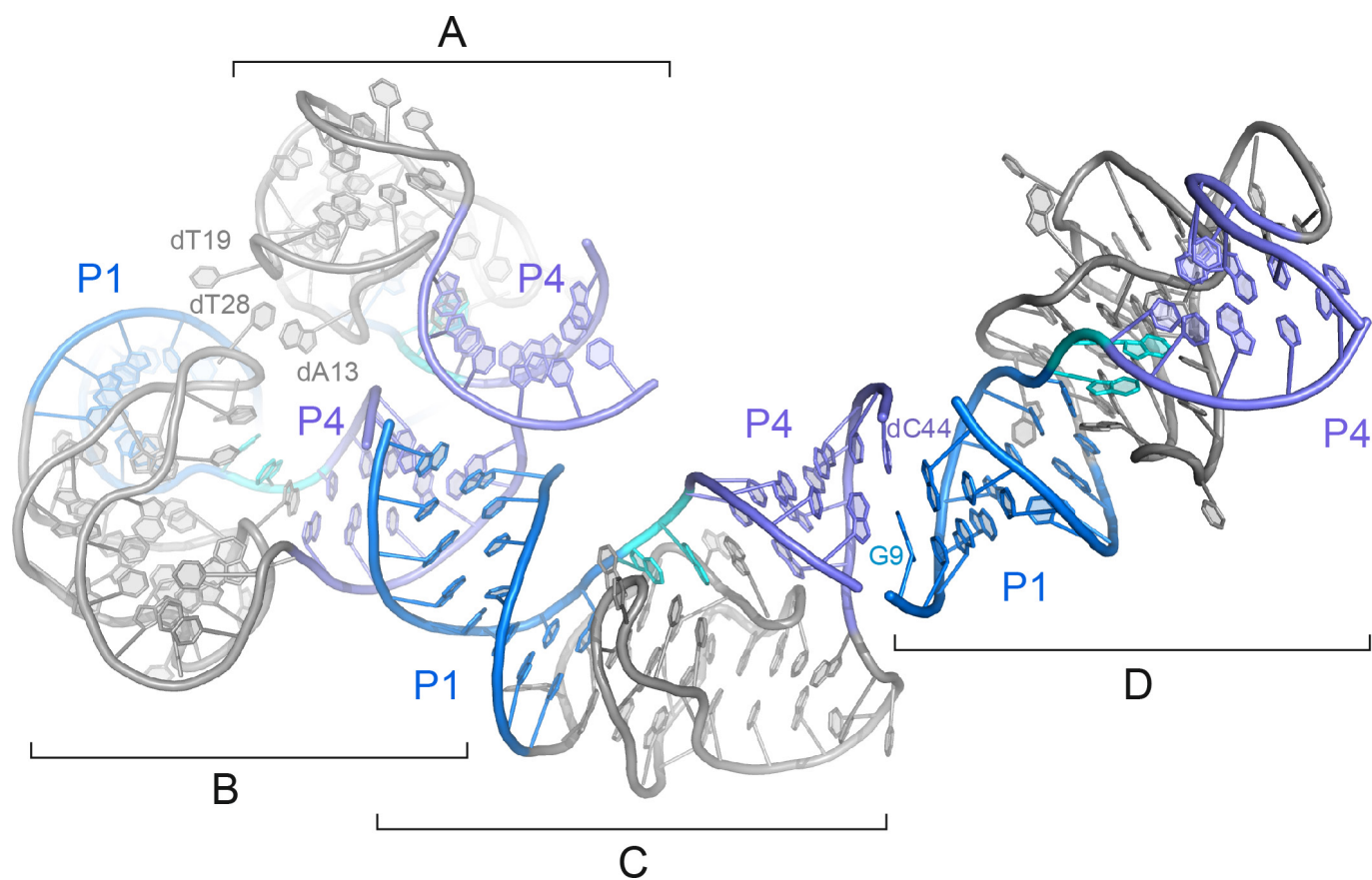
Prior to data collection, crystals were cryoprotected with perfluoropolyether (PFPE) and flash frozen. The crystallographic experiments were performed on the PXII beamline at the Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland. The data sets were processed with XDS³⁷.

Crystallographic analyses. The structure of the complex was determined by single anomalous dispersion (SAD) using a hexammine cobalt (III) chloride derivative. The determination of the heavy atom substructure was done with SHELXD³⁸. Initial SAD phase calculation and automatic model building were carried out with AutoSol wizard³⁹. The resulting model from AutoSol was manually edited with COOT⁴⁰. The edited partial model was used along with the experimentally determined cobalt sites for MR-SAD phase calculation within Phaser⁴¹. The resulting map was solvent flattened using RESOLVE⁴². This map was used for tracing the entire phosphodiester backbone. The RNA strand of the model was built using Rcrane⁴³ and the DNA strand was built from individual nucleotide monophosphates. Structure refinement was carried out with phenix.refine⁴⁴. Analysis of pseudorotation phase angles was performed on the PROSIT webserver (<http://cactus.nci.nih.gov/prosit/>)⁴⁵.

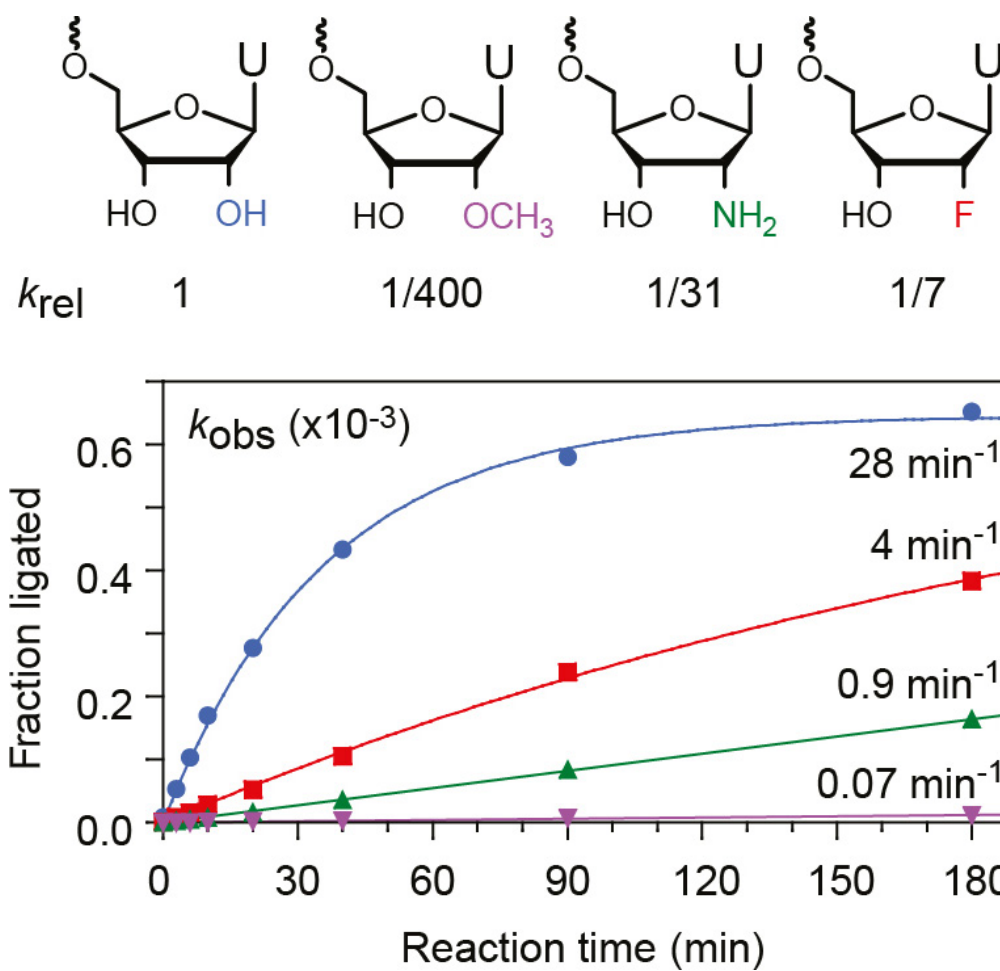
Kinetic assays of DNA-catalysed RNA ligation reactions. Single-turnover assays were performed as described previously^{14,15}. Briefly, 5'-³²P-trace-labelled acceptor

RNA (3 pmol), 9DB1 deoxyribozyme, and 5'-triphosphorylated donor RNA were mixed in a 1:5:10 ratio and annealed by heating to 95 °C for 2 min, followed by slowly cooling for 15 min to room temperature (for RNA/DNA sequences see Extended Data Table 2). The assays were performed in 10 µl volume at 37 °C in buffer containing 50 mM HEPES, pH 7.5, 150 mM NaCl, 2 mM KCl and 20 mM MnCl₂, or 50 mM Tris buffer, pH 8.5, 150 mM NaCl, 2 mM KCl and 80 mM MgCl₂ including 0, 0.1, 1, or 10 mM CdCl₂. Aliquots of 1 µl were removed at desired time points and quenched in stop solution (80% formamide, 1 × TBE, 50 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol). The samples were analysed by PAGE (15% polyacrylamide), and band intensities were quantified by Phosphorimaging. The yield versus time data were fit to (fraction ligated) = $Y(1 - e^{-kt})$, where $k = k_{\text{obs}}$ and Y = final yield. No statistical methods were used to predetermine sample size.

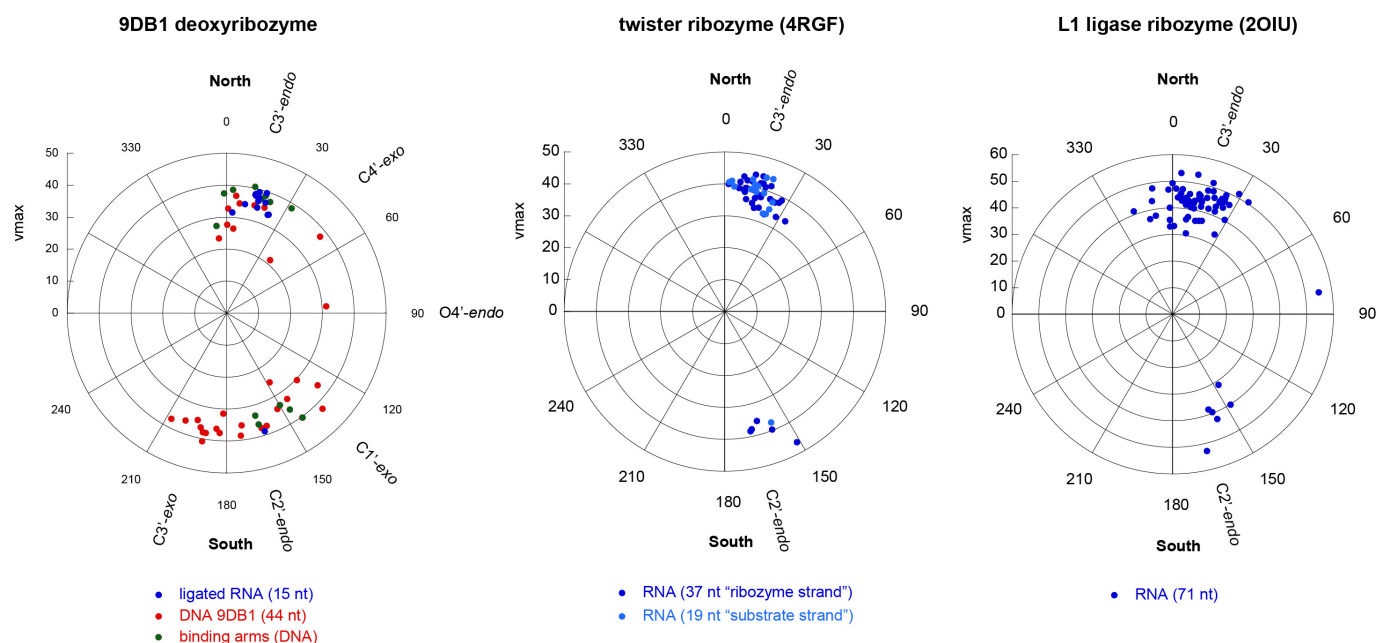
31. Pitsch, S., Weiss, P. A., Jenny, L., Stutz, A. & Wu, X. Reliable chemical synthesis of oligoribonucleotides (RNA) with 2'-O-[(Triisopropylsilyl)oxy]methyl(2'-O-TOM)-protected phosphoramidites. *Helv. Chim. Acta* **84**, 3773–3795 (2001).
32. Kumar, R. K., Olsen, P. & Ravikumar, V. T. An alternative advantageous protocol for efficient synthesis of phosphorothioate oligonucleotides utilizing phenylacetyl disulfide (PADS). *Nucleosides Nucleotides Nucleic Acids* **26**, 181–188 (2007).
33. Frederiksen, J. K. & Piccirilli, J. A. Separation of RNA phosphorothioate oligonucleotides by HPLC. *Methods Enzymol.* **468**, 289–309 (2009).
34. Koch, M. et al. Role of a ribosomal RNA phosphate oxygen during the EF-G-triggered GTP hydrolysis. *Proc. Natl Acad. Sci. USA* **112**, E2561–E2568 (2015).
35. Ludwig, J. & Eckstein, F. Rapid and efficient synthesis of nucleoside 5'-O-(1-thiotriphosphates), 5'-triphosphates and 2',3'-cyclophosphorothioates using 2-chloro-4h-1,3,2-benzodioxaphosphorin-4-one. *J. Org. Chem.* **54**, 631–635 (1989).
36. Goldeck, M., Tuschl, T., Hartmann, G. & Ludwig, J. Efficient solid-phase synthesis of pppRNA by using product-specific labeling. *Angew. Chem. Int. Ed. Engl.* **53**, 4694–4698 (2014).
37. Kabsch, W. X. D. S. *Acta Crystallogr. D* **66**, 125–132 (2010).
38. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D* **66**, 479–485 (2010).
39. Terwilliger, T. C. et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D* **65**, 582–601 (2009).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
41. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
42. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
43. Keating, K. S. & Pyle, A. M. RCrane: semi-automated RNA model building. *Acta Crystallogr. D* **68**, 985–995 (2012).
44. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
45. Sun, G., Voigt, J. H., Filippov, I. V., Marquez, V. E. & Nicklaus, M. C. PROSIT: pseudo-rotational online service and interactive tool, applied to a conformational survey of nucleosides and nucleotides. *J. Chem. Inf. Comput. Sci.* **44**, 1752–1762 (2004).
46. Ren, A. et al. In-line alignment and Mg²⁺ coordination at the cleavage site of the env22 twister ribozyme. *Nat. Commun.* **5**, 5534 (2014).



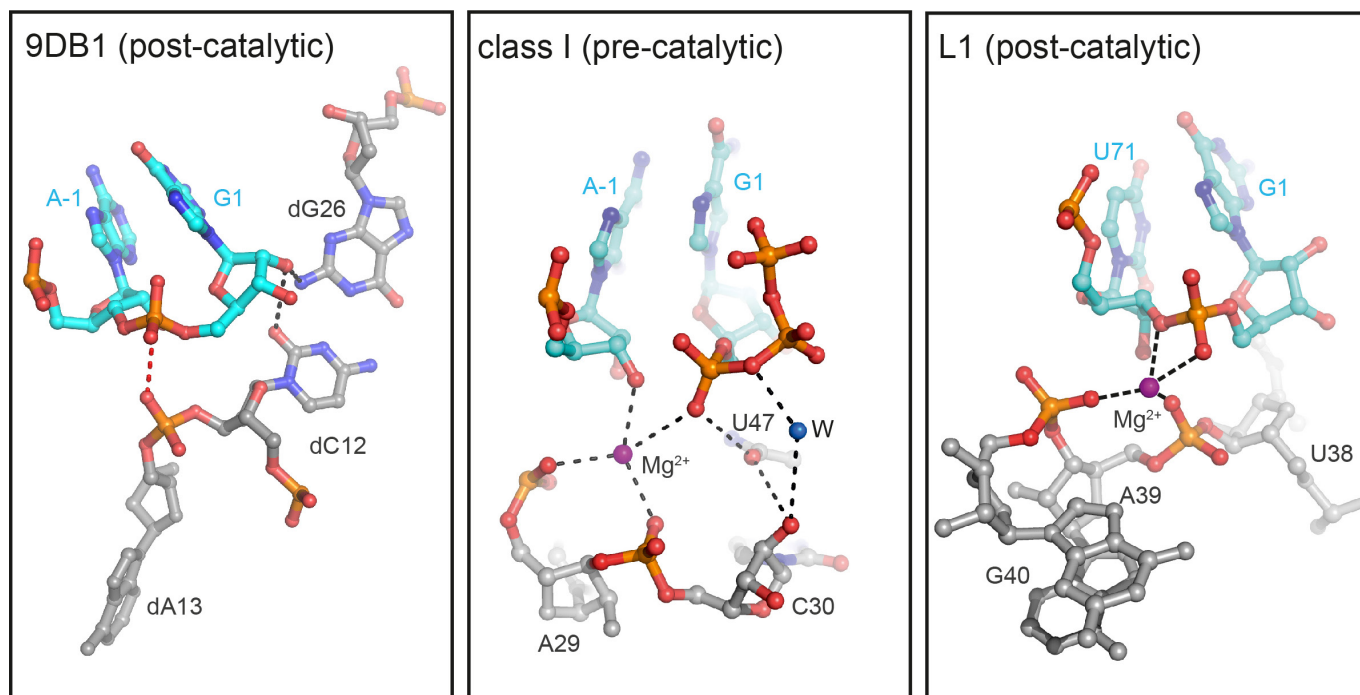
Extended Data Figure 1 | Semi-continuous helix in the crystal lattice. Nucleotides important for the crystal contacts are labelled accordingly.



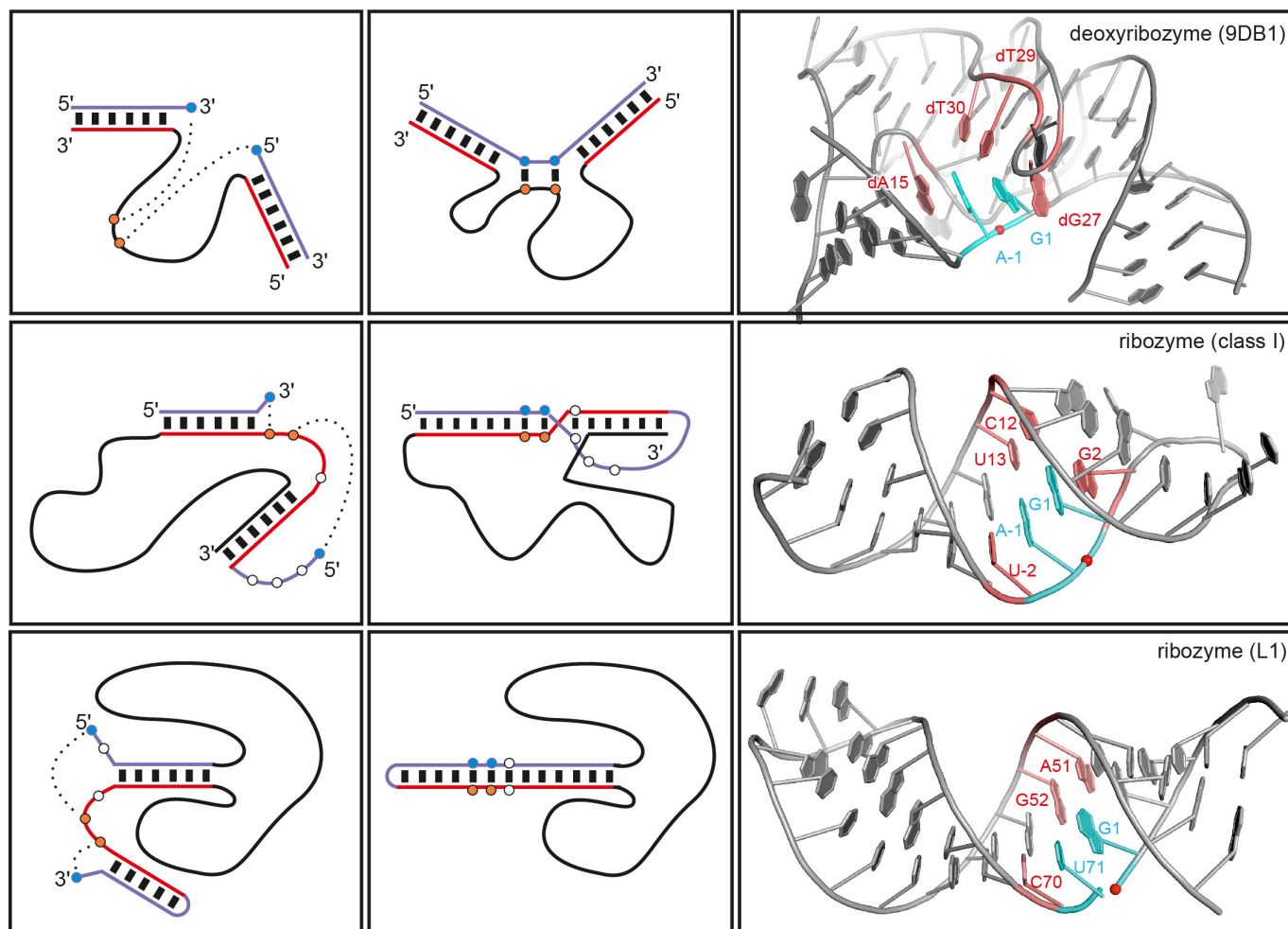
Extended Data Figure 2 | Influence of 2'-modifications on mutant acceptor RNA U-1.



Extended Data Figure 3 | Distribution of pseudorotation phase angles of 9DB1 deoxyribozyme in comparison with twister⁴⁶ and L1 ligase ribozyme³⁰.



Extended Data Figure 4 | Active sites of RNA-ligating nucleic acid enzymes. Left: deoxyribozyme 9DB1; middle: class I ribozyme; right: L1 ligase ribozyme. Figure was generated using Protein Data Bank (PDB) accession numbers 5CKK (9DB1), 3R1L (class I) and 2OIU (L1).



Extended Data Figure 5 | Positioning of the reactive nucleotides in the active centre of ligase deoxyribozyme 9DB1 in comparison to ribozymes. A schematic of the three enzymes is shown before and after the crystal structure determination (left and central columns, respectively). Strands that were deliberately chosen are depicted in red and blue, while the ones that resulted from random selection are shown in

black. Nucleotides at the ligation junction are coloured in cyan, stacking nucleotides are depicted in light red. Nucleotides shown as white-filled circles are indicated for orientation purposes. Recognition between reactive and pairing nucleotides is shown as dashed lines (left column). Structures of product-bound nucleic acid enzymes are compared (PDB accessions 5CKK (9DB1), 3HHN (class I) and 2OIU (L1), right column).

Extended Data Table 1 | Data collection and refinement statistics

	Native	Co-Hex
Data collection		
Space group	P 43 21 2	P 43 21 2
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	83.1, 83.1, 55.9	80.8, 80.8, 55.8
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength	1.0	1.6
Resolution (Å)	41.58 – 2.80 (2.90 – 2.80)	40.42 – 2.98 (3.09 – 2.98)
<i>R</i> _{merge} (%)	5 (>100)	8 (>100)
<i>I</i> / σ <i>I</i>	33.87 (0.43)	25.23 (0.86)
CC _{1/2}	0.99 (0.90)	0.99 (0.52)
Completeness (%)	97 (97)	100 (98)
Redundancy	24.8 (25.8)	26.1 (24.0)
Refinement		
Resolution (Å)	2.80	2.98
No. reflections	5012	4053
<i>R</i> _{work} / <i>R</i> _{free}	0.27/ 0.29	0.25/ 0.31
No. atoms		
Nucleic acid	1232	1232
Ligand/ion	3	6
B-factors		
Nucleic acid	148.07	133.72
Ligand/ion	119.43	132.84
R.m.s deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.66	0.60

One crystal was used for each data set. Highest-resolution shell values are shown in parentheses.

Extended Data Table 2 | Oligonucleotides for kinetic experiments

description	5'-sequence-3'
acceptor RNA 3'-A	GGAAGUCUCAUGUACUA
acceptor RNA 3'-(2'dA)	GGAAGUCUCAUGUACU (dA)
acceptor RNA 3'-(2-OMeA)	GGAAGUCUCAUGUACU (OMeA)
acceptor RNA 3'-(2'-F-A)	GGAAGUCUCAUGUACU (FA)
acceptor RNA 3'-U	GGAAGUCUCAUGUACUU
acceptor RNA 3'-(2'-OMe-U)	GGAAGUCUCAUGUACU (OMeU)
acceptor RNA 3'-(2'-NH ₂ -U)	GGAAGUCUCAUGUACU (NH ₂ U)
acceptor RNA 3'-(2'-F-U)	GGAAGUCUCAUGUACU (FU)
donor RNA1 5'pppG	pppGAUGUUCUAGCGCCGA
donor RNA2 5'pppG	ppp G ACGCUGACCCUGAAGUUCAUCUU
donor RNA 5'pppA	ppp A ACGCUGACCCUGAAGUUCAUCUU
donor RNA 5'pppC	ppp C ACGCUGACCCUGAAGUUCAUCUU
donor RNA 5'pppU	ppp U ACGCUGACCCUGAAGUUCAUCUU
9DB1 DNA original*	CAAGGCGCTAGAACATGGATCATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGACTTCC
9DB1 DNA*,†	GAAC TTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGT T TGCCGTTTAAGTACATGAGACTTCC
9DB1 dT29C	GAAC TTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGT C TGCCGTTTAAGTACATGAGACTTCC
9DB1 dT29A	GAAC TTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGT A TGCCGTTTAAGTACATGAGACTTCC
9DB1 dT29G	GAAC TTCAGGGTCAGCGTGGATCATA CGGTCGGAGGGGT G TGCCGTTTAAGTACATGAGACTTCC
9DB1 DNA (Fig. 4b)	GCGCTAGAACATGGATCATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGAC
9DB1 DNA (ps) (Fig. 4b)‡	GCGCTAGAACATGGATC ps ATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGAC
9DB1 DNA (pCH3) (Fig. 4b)‡	GCGCTAGAACATGGATC (pCH3) ATA CGGTCGGAGGGGTTTGCCGTTTAAGTACATGAGAC
splint for ligation	CCCCGCCGACCGTATGATCCATGTTC

*Catalytic core is underlined.

†dT29 is in bold.

‡Ligation site is indicated with |.

Structures of two distinct conformations of holo–non–ribosomal peptide synthetases

Eric J. Drake^{1,2*}, Bradley R. Miller^{1,2*}, Ce Shi³, Jeffrey T. Tarrasch⁴, Jesse A. Sundlov^{1,2}, C. Leigh Allen¹, Georgios Skiniotis⁴, Courtney C. Aldrich³ & Andrew M. Gulick^{1,2}

Many important natural products are produced by multidomain non-ribosomal peptide synthetases (NRPSs)^{1–4}. During synthesis, intermediates are covalently bound to integrated carrier domains and transported to neighbouring catalytic domains in an assembly line fashion⁵. Understanding the structural basis for catalysis with non-ribosomal peptide synthetases will facilitate bioengineering to create novel products. Here we describe the structures of two different holo–non–ribosomal peptide synthetase modules, each revealing a distinct step in the catalytic cycle. One structure depicts the carrier domain cofactor bound to the peptide bond-forming condensation domain, whereas a second structure captures the installation of the amino acid onto the cofactor within the adenylation domain. These structures demonstrate that a conformational change within the adenylation domain guides transfer of intermediates between domains. Furthermore, one structure shows that the condensation and adenylation domains simultaneously adopt their catalytic conformations, increasing the overall efficiency in a revised structural cycle. These structures and the single-particle electron microscopy analysis demonstrate a highly dynamic domain architecture and provide the foundation for understanding the structural mechanisms that could enable engineering of novel non-ribosomal peptide synthetases.

A non-ribosomal peptide synthetase (NRPS) module incorporates a single residue into a peptide natural product. Each module contains a peptidyl carrier protein (PCP) that is post-translationally modified with a phosphopantetheine cofactor⁶, an adenylation domain that loads the amino-acid substrate onto the PCP cofactor, and a condensation domain that catalyses peptide bond formation. NRPSs then use a carboxy (C)-terminal thioesterase or reductase domain to catalyse product release. Structures of individual domains¹ provide insight into the NRPS structural mechanism. Interestingly, the adenylation domains have been shown to adopt two catalytic conformations⁷. First the adenylate-forming conformation activates the amino-acid substrate using ATP to form an aminoacyl adenylate and pyrophosphate. A C-terminal subdomain then rotates by ~140° to form the thioester-forming conformation that is used to install the amino acid onto the PCP⁷. These two functional states have been observed in structures of the phenylalanine activating adenylation domain of gramicidin synthetase⁸ and the complexes between adenylation and carrier proteins obtained with mechanism-based inhibitors^{9,10}. Once loaded, both the pantetheine and loaded substrate have been shown to interact transiently with the core of the carrier protein^{11,12}. The structure of SrfA-C, the terminal module from surfactin biosynthesis, contains a condensation–adenylation–PCP–thioesterase architecture and is to date the only structure of an intact NRPS module¹³. The condensation and adenylation domains share an extensive interface and were proposed to form the core of the module¹³. Lacking the pantetheine modification, this apo-structure shows the PCP domain directed towards the condensation domain. The other active sites are 40–60 Å from the

pantetheinylation site, indicating that extensive domain rearrangements are required to complete the NRPS catalytic cycle. Movement of the PCP domain, potentially coupled to the adenylation C-terminal subdomain rotation⁷, is necessary for delivery of the peptide intermediates to the different catalytic domains.

We determined structures of two NRPSs with the same architecture as SrfA-C (Extended Data Fig. 1), but with holo-proteins that show functional interactions between the PCP and catalytic domains (Fig. 1). First we present two structures of AB3403 from the human pathogen *Acinetobacter baumannii* (protein annotation ABBFA_003403 in strain AB307-0294) that belongs to an uncharacterized biosynthetic pathway implicated in motility¹⁴, and biofilm¹⁵ and pellicle¹⁶ formation. We describe the structures of holo-AB3403 obtained without ligands and also upon crystallization in the presence of Mg-ATP and glycine, which among the proteinogenic amino acids serves as the best substrate (Extended Data Fig. 2). Second, we present the structure of EntF from *Escherichia coli*, showing the PCP cofactor covalently trapped with a mechanism-based inhibitor to model thioester formation within the adenylation domain. These results provide views of two distinct steps in the NRPS catalytic cycle and demonstrate how the domain rotation within the adenylation domain mediates the delivery of the PCP between the two catalytic domains.

The structures of AB3403 were determined at 2.7 and 2.9 Å resolution (Extended Data Table 1). No prior structure exists of an NRPS condensation domain bound to a ligand; the holo-AB3403 protein shows the pantetheine cofactor residing in the active site (Fig. 2 and Extended Data Fig. 3a). The two lobes of the condensation domain adopt the closed orientation seen recently in the CDA synthetase condensation domain¹⁷. Contacts are made between the pantetheine and the helix running from Glu20 to Leu30, in particular Tyr26 and Ile27, which forms one wall of the tunnel through which the pantetheine approaches the active site (Fig. 2b). Additionally, Tyr37 forms a hydrogen bond with the amide of the cysteamine moiety of the pantetheine cofactor. As the main chain carbonyl of Tyr37 hydrogen bonds to the main chain amide of the catalytic His145, this is a critical interaction to close the two lobes and bring the active histidine into proper position.

Holo-AB3403 therefore illustrates the conformation that is adopted to properly deliver the pantetheine of the PCP to the condensation domain. The PCP is rotated ~30° relative to the orientation of the PCP domain of SrfA-C (Extended Data Fig. 4). The AB3403 PCP interface with the condensation domain is composed of residues from helix α_2 , the helix that follows the pantetheinylation site at Ser1006, and the loops that precede and follow this helix. In particular, residues Phe999 to Tyr1032 face the condensation domain. Leu1007 and Val1010 form a hydrophobic interaction with Leu22 and Ile80 of the condensation domain. The side chain of Lys1011 forms a hydrogen bond with the main chain carbonyl of Gln78. Finally, Val1026, Ala1027, and Ala1030 on the PCP helix α_3 form a hydrophobic interaction with Tyr26 and Leu30. Arg344 of the

¹Hauptman-Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, New York 14203, USA. ²Department of Structural Biology, University at Buffalo, Buffalo, New York 14203, USA.

³Center for Drug Design and Department of Medicinal Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA. ⁴Life Sciences Institute and Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA.

*These authors contributed equally to this work.

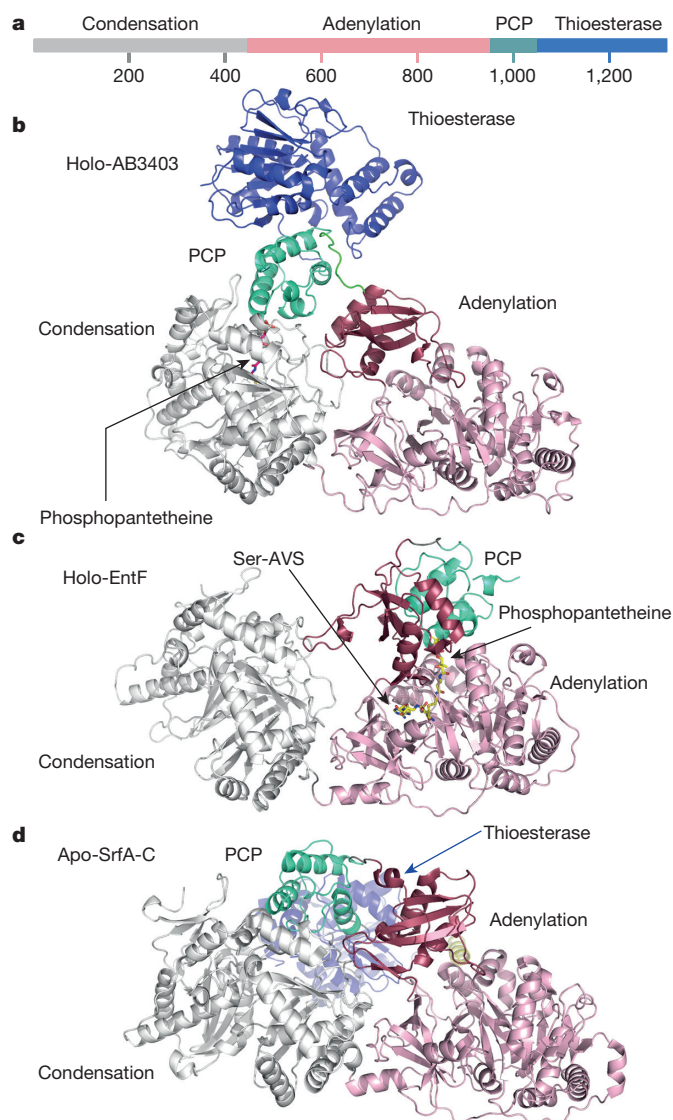


Figure 1 | Ribbon diagrams of complete NRPS modules. **a**, Domain architecture of three structurally characterized termination modules. **b–d**, The protein structures of **(b)** AB3403, **(c)** EntF, and **(d)** SrfA-C are coloured with domains coloured white (condensation), pink and red (adenylation domain N- and C-terminal subdomains), green-cyan (PCP), and blue (thioesterase). The phosphopantetheine moieties of AB3403 and EntF, and inhibitor Ser-AVS, are highlighted.

condensation domain, which is positioned on an insertion compared with SrfA-C, interacts with the phosphate from the cofactor.

The AB3403 adenylation domain (Fig. 2c) is precisely positioned in the adenylate-forming conformation, unlike the adenylation domain of SrfA-C, which is in an open conformation that may be used for substrate binding or release⁵. The lysine of the conserved catalytic A10 motif^{7,18} interacts with a phosphate oxygen from AMP and a carboxylate oxygen with glycine and superimposes with the homologous lysine in the gramicidin synthetase domain. In SrfA-C, the homologous lysine is ~12 Å away.

The thioesterase domain of AB3403 is structurally similar to the homologous domains of both SrfA-C and EntF (Extended Data Fig. 5), the latter of which has been characterized by NMR and crystallography in complex with the upstream PCP domain^{19,20}. Despite the similarities in domain structure, the thioesterase domain of AB3403 is in a markedly different location compared with SrfA-C (Fig. 3a). Interestingly, in this new position the thioesterase domain cradles the back face of the PCP domain. The thioesterase domains of SrfA-C or AB3403 do not make substantial contacts with the other catalytic domains.

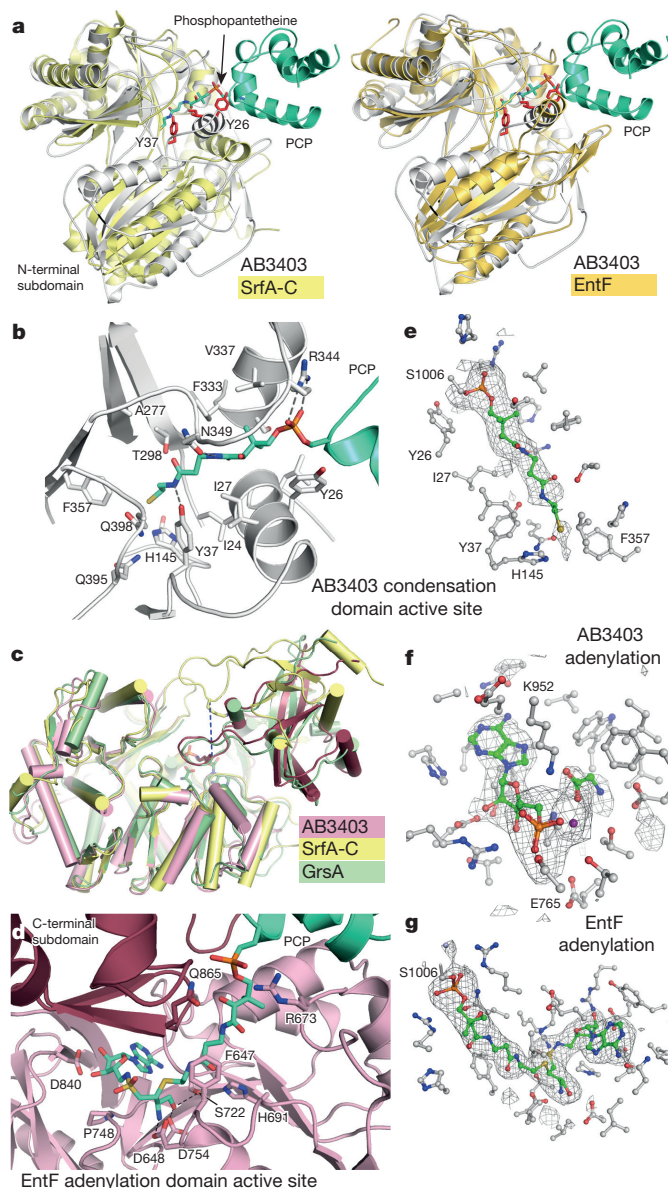


Figure 2 | NRPS domain structures. **a**, The condensation domain of AB3403 (white) was aligned with SrfA-C (yellow) and EntF (orange) on the basis of the condensation C-terminal subdomain. The AB3403 PCP is included. **b**, The AB3403 condensation domain highlights residues that form the hydrophobic tunnel through which the pantetheine passes. **c**, Superposition of adenylation domains of AB3403 (pink and maroon for N- and C-terminal subdomains), SrfA-C (yellow) and gramicidin synthetase, GrsA (cyan), with phenylalanine and AMP molecules of GrsA. The dotted line highlights the alternative position of the catalytic lysines of AB3403 and SrfA-C. **d**, The EntF adenylation domain active site shows a covalent linkage from the pantetheine to the Ser-AVS inhibitor. **e–g**, Electron density calculated with coefficients of the form $F_o - F_c$ generated before inclusion of ligands and contoured at 3σ , are shown for the **(e)** AB3403 condensation, **(f)** AB3403 adenylation, and **(g)** EntF adenylation domains.

We next examined the delivery of the holo-PCP to the adenylation domain in a different NRPS protein. We have previously used targeted mechanism-based inhibitors, harbouring a vinylsulfonamide moiety that traps the thioester-forming reaction²¹ to characterize functional adenylation-PCP di-domain interactions^{9,10}. These inhibitors mimic the native aminoacyl adenylate, but contain a Michael acceptor positioned to react with the pantetheine thiol. EntF crystallized only in the presence of the serine adenosine vinylsulfonamide (Ser-AVS) inhibitor (Fig. 2d and Extended Data Fig. 6) that limits conformational flexibility

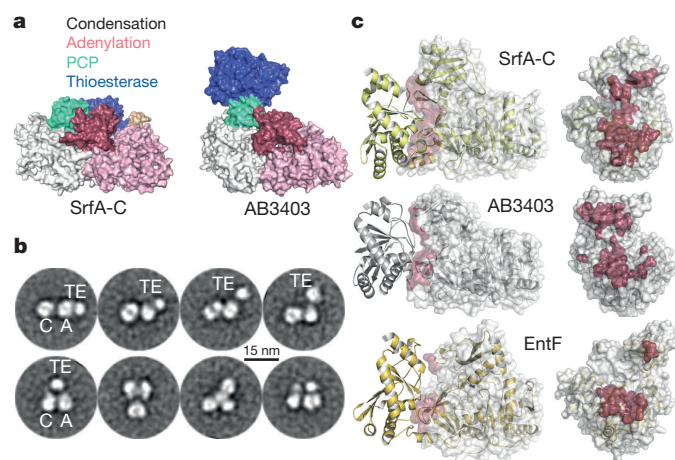


Figure 3 | Conformational dynamics in NRPS modules. **a**, Alternative locations of the thioesterase domain SrfA-C and AB3403. **b**, Representative electron microscopy class averages of EntF. The smaller thioesterase (TE) domain is observed in various positions relative to the condensation (C)–adenylation (A) di-domain. Overall EntF adopts a variety of extended (top) to compact (bottom) conformations. **c**, The interface between the condensation C-terminal subdomain and the adenylation domain is shown for SrfA-C, AB3403, and EntF. The adenylation surface is shown in white, highlighting in red the regions that interact with the condensation domain. The right panel shows this interface, rotated by 90° around the y axis, with the condensation domain omitted for clarity.

to promote crystallization. Crystals of the EntF protein diffract to 2.8 Å (Extended Data Table 2). No electron density was observed for the thioesterase domain although the intact protein was present in the crystal lattice (Extended Data Fig. 7).

The condensation domain of EntF is similar to the closed AB3403 conformation (Fig. 2a). The adenylation domain adopts the catalytic thioester-forming conformation of prior adenylation-PCP proteins^{9,10}, demonstrating that the conformation is compatible with a full NRPS module. The active site of the EntF adenylation domain identifies conserved residues (Fig. 2d) that have been shown to play important catalytic roles in other members of this enzyme superfamily⁷. Arg863 interacts with the cofactor phosphate, while Gly864 and Gln865 form one wall of the pantetheine tunnel. Interactions with the nucleotide occur between Asp840 and the ribose hydroxyls, and between Tyr746 and Tyr852 and the adenine ring. The inhibitor serine binds in the binding pocket formed by Asp648, Ser722, and Asp754 (Fig. 2d).

The lack of density for the thioesterase domain in EntF suggested multiple conformations in the crystal lattice. This is not surprising given the limited interactions in SrfA-C and AB3403 between the thioesterase domains and the other domains. To assess thioesterase conformational mobility, we examined EntF by negative-stain electron microscopy followed by classification and averaging of single-particle projections (Extended Data Fig. 8). The class averages revealed primarily a tri-lobed density with two neighbouring globular densities of similar size attributed to the condensation and adenylation domains and a smaller lobe attributed to the thioesterase domain (Fig. 3b). The positioning of the thioesterase domain assumes a surprisingly wide range of distances and angles relative to the other domains.

The large interface of the SrfA-C condensation and adenylation domains¹³ suggested they constitute a catalytic platform, upon which the other domains move. We therefore compared the interfaces of the three NRPS modules (Fig. 3c). The interface in AB3403 is 1,023 Å², comparable in size to the 1,097 Å² interface of SrfA-C. In contrast, the interface in EntF is only 780 Å², resulting from the rotation of the adenylation C-terminal subdomain to the thioester-forming conformation.

Additionally, the conformation of the interface is not conserved between all three proteins. Alignment of the structures on the basis of the amino (N)-terminal subdomains of the adenylation domain shows

that the condensation domain of both AB3403 and EntF differ slightly from each other and more significantly from SrfA-C. In AB3403 and EntF, the condensation domains are rotated by ~25° relative to the adenylation domains. Furthermore, the EntF condensation domain is shifted closer towards the adenylation domain. Structural comparisons suggest that this alternative conformation in EntF may not be compatible with the adenylate-forming conformation. The three different condensation–adenylation domain conformations, the adenylate-forming incompatibility seen in EntF, and the multiple extended and compact conformations seen in the electron microscopy data suggest that the condensation–adenylation domain platform may be more dynamic than previously proposed¹³.

The new structures confirm the hypothesis⁷ that the adenylation domain conformational change is a structural mechanism to guide the PCP between active sites in the context of complete NRPS modules. The rotation of the adenylation domain C-terminal subdomain from the adenylate-forming conformation in AB3403 to the thioester-forming conformation of EntF delivers the PCP into the adenylation domain for loading. The recent structure of loaded holo-PCP has shown the interaction of the substrate with the PCP core which may help to promote release of the substrate from the adenylation domain¹¹. This interaction also alters the surface electrostatic potential of regions that interact with the neighbouring catalytic domains, including $\alpha 2$ and $\alpha 3$, and may influence the PCP delivery to neighbouring catalytic domains. Finally, this transfer is further assisted by the linker region that joins the adenylation C-terminal subdomain with the PCP domain, which includes important contacts that are preserved in the adenylate- and thioester-forming conformations²², as well as the open conformation of SrfA-C.

The basic NRPS catalytic cycle requires that the PCP visits three adjacent catalytic domains in a coordinated manner. The two catalytic conformations of the adenylation domain⁷ require that the full cycle has four catalytic structural states (Fig. 4). Specifically, (I) the adenylation domain catalyses amino-acid adenylation, (II) the PCP is

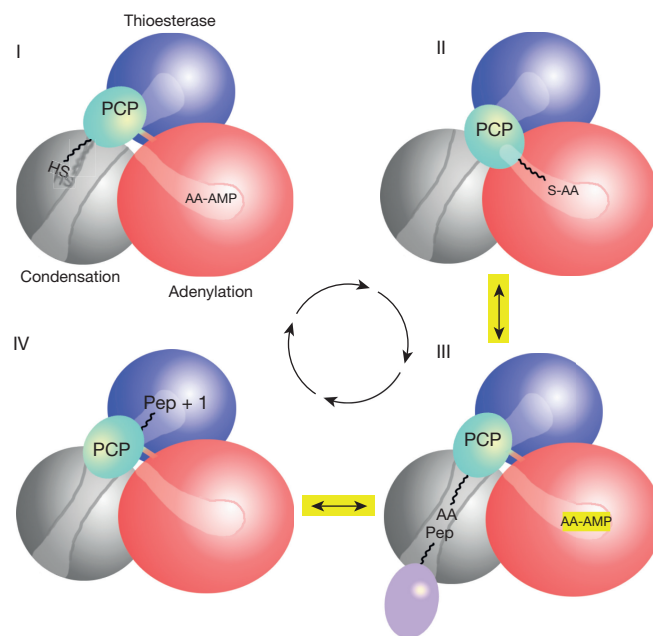


Figure 4 | Dynamics of the NRPS cycle. The four-stage catalytic cycle of an NRPS module. The pantetheine cofactor is represented by the wavy line with a terminal thiol, SH. The aminoacyl adenylate intermediate is represented by AA-AMP. The thioester between the amino acid and the cofactor is shown as S-AA. Finally, the peptide bound to the upstream carrier protein (purple) is abbreviated Pep. Following the condensation reaction, the peptide is extended by one amino acid (Pep + 1) and presented to the thioesterase domain. The revised NRPS structural cycle is highlighted in yellow showing that only three structural states are required.

delivered to the adenylation domain for thioester-formation to load the PCP, (III) the PCP is delivered to the condensation domain to receive the upstream peptide, and finally (IV) the peptide is delivered to a downstream condensation, thioesterase, or reductase domain for release.

Our results show that states I and III are identical and only three distinct conformations are required to accommodate the four catalytic states of the NRPS cycle (Fig. 4, yellow). The protein first adopts an adenylate-forming conformation, seen in AB3403, state III, to catalyse amino-acid adenylation. Through the domain rotation of the adenylation C-terminal subdomain, the PCP is delivered to the adenylation domain to load the pantetheine cofactor, as seen in the crystal structure of EntF, state II. Return of the PCP to the condensation domain delivers the loaded PCP for receipt of the upstream peptide, state III. Critically, as seen in AB3403, the adenylation domain can activate a second amino acid to prime the system for another cycle. The ability to simultaneously catalyse peptide bond formation and amino-acid adenylation at two active sites significantly increases the overall catalytic efficiency and throughput of the NRPS module. Finally, although no structure exists of a full NRPS module with the PCP directed into the thioesterase or other downstream domain in state IV, the structure of AB3403 also offers a new view of the thioesterase domain and suggests the peptide-loaded PCP could be delivered to the downstream thioesterase domain through a simple rotation.

The modular architecture of NRPSs as well as their capacity to catalyse unusual chemistry^{23,24} offer the potential for generating novel products through engineering enzyme activity and the combination of heterologous domains. These efforts have been limited by deficiencies in our understanding of the functional interactions between domains and within active sites. The new views of two essential catalytic states in the NRPS cycle, an appreciation of the greater dynamics of NRPS systems, and the structures of holo-NRPS proteins with relevant ligands will provide the necessary insights to guide these engineering efforts. In addition, these studies complement the recent visualization of modular polyketide synthases by cryo-electron microscopy²⁵ to set the stage for investigations of the structural foundation of even larger, multi-modular biosynthetic proteins.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 June; accepted 29 October 2015.

- Koglin, A. & Walsh, C. T. Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat. Prod. Rep.* **26**, 987–1000 (2009).
- Fischbach, M. A. & Walsh, C. T. Antibiotics for emerging pathogens. *Science* **325**, 1089–1093 (2009).
- Walsh, C. T. The chemical versatility of natural-product assembly lines. *Acc. Chem. Res.* **41**, 4–10 (2008).
- Walsh, C. T. & Fischbach, M. A. Natural products version 2.0: connecting genes to molecules. *J. Am. Chem. Soc.* **132**, 2469–2493 (2010).
- Strieker, M., Tanović, A. & Marahiel, M. A. Nonribosomal peptide synthetases: structures and dynamics. *Curr. Opin. Struct. Biol.* **20**, 234–240 (2010).
- Mercer, A. C. & Burkart, M. D. The ubiquitous carrier protein—a window to metabolite biosynthesis. *Nat. Prod. Rep.* **24**, 750–773 (2007).
- Gulick, A. M. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* **4**, 811–827 (2009).
- Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174–4183 (1997).
- Mitchell, C. A., Shi, C., Aldrich, C. C. & Gulick, A. M. Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains. *Biochemistry* **51**, 3252–3263 (2012).

- Sundlov, J. A., Shi, C., Wilson, D. J., Aldrich, C. C. & Gulick, A. M. Structural and functional investigation of the intermolecular interaction between NRPS adenylation and carrier protein domains. *Chem. Biol.* **19**, 188–198 (2012).
- Goodrich, A. C., Harden, B. J. & Frueh, D. P. Solution structure of a nonribosomal peptide synthetase carrier protein loaded with its substrate reveals transient, well-defined contacts. *J. Am. Chem. Soc.* **137**, 12100–12109 (2015).
- Jaremko, M. J., Lee, D. J., Opella, S. J. & Burkart, M. D. Structure and substrate sequestration in the pyoluteorin type II peptidyl carrier protein PltL. *J. Am. Chem. Soc.* **137**, 11546–11549 (2015).
- Tanovic, A., Samel, S. A., Essen, L. O. & Marahiel, M. A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **321**, 659–663 (2008).
- Clemmer, K. M., Bonomo, R. A. & Rather, P. N. Genetic analysis of surface motility in *Acinetobacter baumannii*. *Microbiology* **157**, 2534–2544 (2011).
- Rumbo-Feal, S. *et al.* Whole transcriptome analysis of *Acinetobacter baumannii* assessed by RNA-sequencing reveals different mRNA expression profiles in biofilm compared to planktonic cells. *PLoS One* **8**, e72968 (2013).
- Giles, S. K., Stroehrer, U. H., Eijkelkamp, B. A. & Brown, M. H. Identification of genes essential for pellicle formation in *Acinetobacter baumannii*. *BMC Microbiol.* **15**, 116 (2015).
- Bloudoff, K., Rodionov, D. & Schmeing, T. M. Crystal structures of the first condensation domain of CDA synthetase suggest conformational changes during the synthetic cycle of nonribosomal peptide synthetases. *J. Mol. Biol.* **425**, 3137–3150 (2013).
- Marahiel, M. A., Stachelhaus, T. & Mootz, H. D. Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651–2674 (1997).
- Frueh, D. P. *et al.* Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature* **454**, 903–906 (2008).
- Liu, Y., Zheng, T. & Bruner, S. D. Structural basis for phosphopantetheinyl carrier domain interactions in the terminal module of nonribosomal peptide synthetases. *Chem. Biol.* **18**, 1482–1488 (2011).
- Qiao, C., Wilson, D. J., Bennett, E. M. & Aldrich, C. C. A mechanism-based aryl carrier protein/thiolation domain affinity probe. *J. Am. Chem. Soc.* **129**, 6350–6351 (2007).
- Miller, B. R., Sundlov, J. A., Drake, E. J., Makin, T. A. & Gulick, A. M. Analysis of the linker region joining the adenylation and carrier protein domains of the modular nonribosomal peptide synthetases. *Proteins* **82**, 2691–2702 (2014).
- Gaudelli, N. M., Long, D. H. & Townsend, C. A. β -Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis. *Nature* **520**, 383–387 (2015).
- Maruyama, C. *et al.* A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. *Nature Chem. Biol.* **8**, 791–797 (2012).
- Dutta, S. *et al.* Structure of a modular polyketide synthase. *Nature* **510**, 512–517 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Sanishvili for assistance with data collection. This work was funded in part by National Institutes of Health GM-068440 (to A.M.G.) and GM-115601 (to G.S.), and Award W81XWH-11-2-0218 from the Telemedicine and Advanced Technology Research Center of the US Army Medical Research and Materiel Command (A.M.G.). Data were collected at the GM/CA beamline of the Advanced Photon Source, which is funded by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006) under Department of Energy contract number DE-AC02-06CH11357 to A.P.S. A Stafford Fellowship (to B.R.M.) and support from the Hauptman-Woodward Institute is acknowledged.

Author Contributions C.L.A. characterized activity of and initially crystallized AB3403. J.A.S. initially crystallized EntF. E.J.D. and B.R.M. optimized crystal, and solved and refined the models of AB3403 and EntF, respectively. C.S. and C.C.A. designed and synthesized the mechanism-based inhibitor. J.T.T. and G.S. performed and analysed the single-particle electron microscopy. A.M.G., E.J.D., B.R.M., G.S., J.T.T., C.C.A., and C.S. analysed the results and wrote the manuscript. All authors saw and approved the manuscript.

Author Information The coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession numbers 4ZXH (holo-AB3403), 4ZXI (holo-AB3403 bound to AMP and glycine), and 4ZXJ (holo-EntF bound to Ser-AVS). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.M.G. (gulick@hwi.buffalo.edu).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Expression, purification, and crystallization of AB3403. The human pathogen *A. baumannii* contains an uncharacterized NRPS cluster that has been implicated in motility and biofilm formation; the product of this operon is unknown. This operon contains eight genes. In strain AB307-0294 (ref. 26), from which the NRPS gene was cloned, this operon consists of genes ABBFA_003399 to ABBFA_003406. In the more commonly used ATCC17978 strain, the same genes are encoded by A1S_0119 to A1S_0112. The ABBFA_003403 (designated AB3403 herein) protein sequence is available at GenBank under accession number ACJ56070.1.

The gene encoding AB3403 was PCR-amplified from AB307-0294 genomic DNA²⁶ (courtesy of T. A. Russo). The amplified fragment was cloned into the pET15b-TEV expression vector²⁷ and confirmed by DNA sequencing. The vector provides a His₅-tag, linker, and tobacco etch virus (TEV) protease recognition site that, upon treatment with TEV protease, yields a final recombinant product with glycine and histidine preceding the initial methionine residue.

The AB3403 pET15b-TEV construct was transformed into *E. coli* (BL21-DE3) cells. Transformed cells were grown in LB media to an absorbance at 600 nm ($A_{600\text{ nm}}$) of 0.6 at 37 °C. Protein expression was induced by addition of 0.5 mM isopropyl- β -D-thiogalactoside (IPTG) and cells were incubated overnight at 16 °C. Cells were harvested by centrifugation, flash-frozen in liquid nitrogen, and stored at -80 °C. Selenomethionine-labelled protein was generated in M9 minimal media using a metabolic inhibition protocol²⁸. All purification steps were identical to the native protein.

For purification, cells were resuspended in a buffer containing 50 mM HEPES (pH 7.5), 250 mM NaCl, 10 mM imidazole, 0.2 mM TCEP. Cells were lysed by mechanical disruption (Branson Sonifier) and the resulting lysate was clarified by centrifugation at 235,000g for 45 min. The cell lysate was passed over a His-trap (GE Healthcare) immobilized metal ion affinity column and washed with lysis buffer containing 50 mM imidazole. Bound proteins were eluted with the same buffer containing 300 mM imidazole. The protein was incubated with TEV protease and dialysed against a TEV cleavage buffer (50 mM HEPES (pH 8.0), 250 mM NaCl, 0.2 mM TCEP, and 0.5 mM EDTA) for 16 h at 4 °C. This partly purified protein was then phosphopantetheinylated by incubation with His₆-tagged non-specific phosphopantetheinyl transferase Sfp (10 nM), 12.5 mM MgCl₂, and 1 mM CoA for 60 min at 20 °C. The clarified protein was then passed over the His-trap column a second time to remove uncleaved protein, the TEV protease, Sfp, and other contaminating proteins. The holo-AB3403 protein in the column flow-through was pooled, dialysed against a size exclusion buffer containing 50 mM HEPES (pH 7.5), 150 mM NaCl, 0.2 mM TCEP, and further purified by gel filtration (Superdex200). Protein concentration was assessed after dialysis against a crystallization buffer (25 mM HEPES (pH 7.5), 50 mM NaCl, 0.2 mM TCEP) using an extinction coefficient at 280 nm of 157,570 M⁻¹ cm⁻¹.

Crystallization conditions for holo-AB3403 were initially identified from a sparse matrix screen at 20 °C. Final crystals for native and SeMet-labelled holo-AB3403 were grown at 14 °C by hanging-drop vapour diffusion against 0.75–0.95 M potassium citrate, 0.01–0.025 M glycine, and 0.05 M bis-tris propane (BTP) (pH 8.0). Highest-quality native crystals were obtained using a protein concentration of 5.5 mg ml⁻¹ with a protein to cocktail ratio of 1.5:1. SeMet protein was crystallized in the same manner with a protein concentration of 7.5 mg ml⁻¹ and 1:1 protein to cocktail ratio. To obtain crystals in the presence of ligands, the protein was pre-incubated for 45 min at 4 °C with 2 mM MgCl₂, and 1.5-fold molar excess of ATP and glycine.

Structure determination of AB3403. Crystals of holo-AB3403 were cryoprotected by stages using either ethylene glycol or potassium citrate for native and SeMet protein, respectively. The native protein crystals were cryo-protected with cocktails containing 1.0 M potassium citrate, 0.3 M glycine, 0.05 M BTP (pH 8.0), and increasing (8, 16, and 24%) v/v ethylene glycol. The SeMet-labelled protein was cryo-protected with cocktails containing 0.3 M glycine, 0.05 M BTP (pH 8.0) and increasing (1.0, 1.2, 1.4, and 1.6 M) potassium citrate. Crystals derived from protein co-crystallized with ligands included the same concentration of MgCl₂, ATP, and glycine in the cryo-protectant cocktails.

Diffraction data were collected on APS beamline 23-IDB. The native data (2.7 Å) were collected using a multi-crystal, multi-data set strategy using two crystals. A complete low-resolution scan was taken for one crystal followed by a higher-resolution scan of the best diffracting crystal. A high-resolution region of the second crystal was combined with the two scans from the first crystal. The optimal regions were identified with the JBLU-ICE software at the GM/CA beamline. A single peak wavelength data set (3.35 Å) was collected for SeMet-labelled protein. The liganded protein data set was collected with a single crystal.

Diffraction data were indexed, merged, and scaled using iMOSFLM²⁹ in space group $P4_32_12$. Structure determination was performed with PHENIX³⁰ using a combination of experimental single-wavelength anomalous diffraction (SAD) phasing and phased molecular replacement. A partial molecular replacement solution was positioned through PHASER with a sculpted (PHENIX sculptor) model derived from PheA (PDB accession number 1AMU)⁸ and CytC1 (PDB 3VNR). Using this partial molecular replacement model, the selenium sites were identified with the SAD data from SeMet-labelled crystals. An initial model was produced with PHENIX Autobuild that contained ~65% of the protein molecule, spread across multiple symmetry related molecules. This model was combined into a single protein chain, built and refined iteratively against native data using ARP-WARP³¹, COOT³², and PHENIX refine.

The final refinements were performed with translation-libration-screw-rotation (TLS) parameterization³³ with groups consisting of residues 1:191, 191:445, 446:480, 481:862, 863:959, 960:973, 974:1044, and 1054:1318, roughly defining the NRPS domain (or subdomain) boundaries. The protein is complete from residues Asn2 to Pro1319 with two small disordered loops in the adenylation domain at Asn500–Asp501 and Gly627–Gly630. The latter loop is part of the conserved serine/threonine- and glycine-rich P-loop that is involved in binding the triphosphate of the nucleotide⁷. Additionally, the condensation domain contains electron density for a diacylglycerol lipid molecule that co-purified with the protein and potentially derived from the bacterial membrane during cell disruption. Diffraction and refinement statistics are presented in Extended Data Table 1. Experimental electron densities of the ligands of both structures are presented in stereo format in Extended Data Fig. 3.

Purification of EntF. The enterobactin biosynthetic cluster of *E. coli* has been used as a model system in many studies. The full-length EntF, containing the condensation, adenylation, PCP, thioesterase domain architecture, loads serine onto the PCP domain. The condensation domain then recognizes the external carrier protein EntB that has been loaded with 2,3-dihydroxybenzoate (DHB) by the activity of the freestanding adenylation domain EntE. The DHB-serine amide is then transferred to the thioesterase domain while two additional cycles of synthesis complete the enterobactin trilactone.

The EntF protein used in this study (GenBank P11454) was described previously^{22,34}. The *entf* gene was PCR amplified from *E. coli* JM109 and cloned into a pET15-TEV vector with a N-terminal 5× His-tag and a TEV protease cleavage site²². The *entf* vector was transformed into *E. coli* (BL21-DE3) cells for protein expression. Cells were grown in lysogeny broth (LB) media to $A_{600\text{ nm}} = 0.6$ at 37 °C before protein induction with 1 mM IPTG. Cells were grown overnight at 16 °C and collected by centrifugation. The cell pellets were flash frozen in liquid nitrogen. Selenomethionine-labelled EntF was expressed in M9 minimal media as described²⁸.

For purification both of native and of SeMet-labelled protein, cells were resuspended in lysis buffer containing 50 mM Tris-HCl pH 7.5, 400 mM NaCl, 0.2 mM TCEP, 10% glycerol, and 10 mM imidazole. Cells were lysed via sonication and centrifuged at 235,000g for 45 min. Initial purification was achieved with a His-trap immobilized metal ion affinity column. Protein was eluted using lysis buffer with 300 mM imidazole. EntF was incubated with TEV protease overnight at 4 °C in a cleavage buffer containing 50 mM Tris pH 7.5, 400 mM NaCl, 0.2 mM TCEP, 10% glycerol, and 0.5 mM EDTA. Although expressed in *E. coli*, phosphopantetheinylation was assured by the addition of 10 nM Sfp, 1 mM CoA, and 12.5 mM MgCl₂. The reaction was incubated at room temperature (22 °C) for 1–2 h. The holo-EntF was run over an immobilized metal ion affinity column once more to remove uncleaved protein along with Sfp. A final polishing step was performed with a Superdex 200 16/600 column in a final dialysis buffer containing 50 mM EPPS pH 8.0, 150 mM NaCl, 0.2 mM TCEP, 1 mM MgCl₂, and 10% glycerol. Before crystallization, the Ser-AVS inhibitor was added at a concentration four times that of EntF and allowed to incubate for 2–4 h at room temperature.

For electron microscopy, native EntF was purified as above with the exception that a minimal dialysis buffer was used, which contained 50 mM EPPS pH 8.0, 100 mM NaCl, and 0.2 mM TCEP. No inhibitor was added.

Crystal conditions for the Ser-AVS inhibited EntF were first identified using the Hauptman-Woodward high-throughput screen³⁵. Large diffraction-quality native and SeMet crystals were grown using hanging drop vapour diffusion at 20 °C. A crystallization cocktail, consisting of 100 mM BTP pH 7.5, 125–150 mM MgCl₂, and 22–28% PEG 4000, was diluted 1:1 with the final dialysis buffer. The hanging drops then combined protein at 30 mg ml⁻¹ and the undiluted crystallization cocktail at a ratio of 1:2. This 'batch mimic' limited the differences between the drop and reservoir and has been successful with other protein samples in our laboratory³⁶.

Structure determination of EntF. Native EntF crystals were cryoprotected by that addition of 2,3-butanediol directly to the crystallization drop to a final concentration of ~10%. SeMet crystals were cryoprotected similarly except with

glycerol to a final concentration of ~20%. Diffraction data were collected on APS beamline 23-IDB using the rastering option to find the optimal spots on both the native the SeMet crystals. Diffraction data were indexed, merged, and scaled using iMOSFLM²⁹ in space group $P4_32_12$. Structure determination for the SeMet inflection data was performed in PHENIX³⁰ using a PhaserEP MR-SAD with a partial molecular replacement solution that was obtained using a sculpted model (generated with PHENIX sculptor) derived from the *Pseudomonas aeruginosa* bidomain adenylation-PCP protein PA1221 (PDB 4DG9)⁹. Automated model building with BUCCANEER was used to build ~65% of the structure³⁷. This partial model from the SeMet data was used as a molecular replacement model for the native data, and the remaining portion of the protein was built by hand (excluding the thioesterase domain, which was unresolved and constitutes about 19%). This model was built and refined iteratively using COOT³² and PHENIX refine. TLS refinement³³ was used in final stages with groups consisting of residues 5:186, 187:429, 430:444, 445:857, 858:964, 965:971, and 972:1045.

The final model showed density for the condensation, adenylation, and PCP domains of EntF; no density was observed for the thioesterase domain. Diffraction and refinement statistics are presented in Extended Data Table 2.

In general, the overall quality of the density was weaker for the N-terminal subdomain of the condensation domain, residues 1–186, probably reflecting the higher mobility of this region of the protein. The average B-factors for different regions of the protein (Extended Data Table 2) support this conclusion.

Negative-stain electron microscopy analysis of EntF. EntF, purified as described above, was prepared for electron microscopy using the conventional negative staining protocol³⁸, and imaged at room temperature with a Tecnai T12 electron microscope operated at 120 kV using low-dose procedures. Images were recorded at a magnification of $\times 71,138$ and a defocus value of $\sim 1.5 \mu\text{m}$ on a Gatan US4000 CCD camera. All images were binned (2 pixels \times 2 pixels) to obtain a pixel size of 4.16 Å on the specimen level. Particles were manually excised using e2boxer (part of the EMAN 2 software suite)³⁹. Two-dimensional reference-free alignment and classification of particle projections was performed using ISAC⁴⁰. A total of 17,431 projections of EntF were subjected to ISAC, producing 133 classes consistent over two-way matching and accounting for 5,344 particle projections (Extended Data Fig. 8B).

Synthesis of serine adenosine vinylsulfonamide. Ser-AVS was synthesized using the protocol summarized in (Extended Data Fig. 6). All reactions were performed under an inert atmosphere of dry Ar in oven-dried (150 °C) glassware. ¹H and ¹³C NMR spectra were recorded on a Varian 600 MHz spectrometer. Proton chemical shifts are reported in parts per million from an internal standard of residual chloroform (7.26 p.p.m.) or methanol (3.31 p.p.m.), and carbon chemical shifts are reported using an internal standard of residual chloroform (77.3 p.p.m.) or methanol (49.1 p.p.m.). Proton chemical data are reported as follows: chemical shift, multiplicity (s, singlet; d, doublet; t, triplet; m, multiplet; br, broad), integration, coupling constant. High-resolution mass spectra were obtained on an Agilent TOF II time of flight/mass spectrometry (TOF/MS) instrument equipped with either an ESI or APCI interface. Thin-layer chromatography (TLC) analyses were performed on TLC silica gel 60F254 from EMD Chemical, and were visualized with ultraviolet light or 10% PMA solution. Purifications were performed by flash chromatography on silica gel (Dynamic Adsorbents, 60A).

Materials. Chemicals, reagents, and solvents were purchased from Sigma Aldrich, Chem-Impex, or Acros Organic Fischer, and were used as received. An anhydrous solvent-dispensing system (J. C. Meyer) using two packed columns of neutral alumina was used for drying tetrahydrofuran (THF), Et₂O, while two packed columns of molecular sieves were used to dry DMF and the solvents were dispensed under argon. Compound 1 was purchased from Chem-Impex and used as received. Compounds 2 (ref. 41) and 4 (ref. 10) were synthesized according to the reported procedures.

tert-Butyl (*R,E*)-4-(2-(*N*-(*tert*-butoxycarbonyl)sulfamoyl)vinyl)-2,2-dimethylloxazolidine-3-carboxylate (3). To a solution of *tert*-butyl (2) (395 mg, 1.0 mmol, 2.0 equiv) in 1:3 DMF–THF (4 ml) at –78 °C, was added a 1 M solution of LiHMDS in THF (2.0 ml, 4.0 equiv) dropwise over 15 min and the solution was stirred at –78 °C for an additional 15 min. Next, Garner's aldehyde (1) (115 mg, 0.5 mmol, 1.0 equiv) in THF (1 ml) was added to the reaction over 15 min. The solution was gradually warmed to 25 °C and stirred for 15 h. The solvent was removed *in vacuo* and the mixture was taken up in H₂O (30 ml). The pH was adjusted to 3–4 with 1 N aqueous HCl, then was extracted with ethyl acetate (EtOAc) (3 \times 20 ml). The combined organic layers were washed with H₂O (30 ml), saturated aqueous NaCl (30 ml), dried (MgSO₄), and concentrated. Purification by flash chromatography (10% EtOAc–hexane to 50% EtOAc–hexanes) afforded the title compound 3 as colourless oil (150 mg, 74%); retardation factor (*R*_f) = 0.50 (50:50 EtOAc–hexanes); [α] +0.9 (c 0.02, CH₂Cl₂); ¹H NMR (600 MHz, CD₃OD) ¹H NMR (600 MHz, CD₃OD) δ 1.45 (s, 3H), 1.48

(m, 9H), 1.51 (s, 9H), 1.60 (s, 3H), 3.83–3.85 (m, 1H), 4.15 (dd, *J* = 12.0, 6.0 Hz, 1H), 4.56–4.58 (m, 1H), 6.64 (d, *J* = 18 Hz, 1H), 6.77–6.81 (m, 1H); ¹³C NMR (150 MHz, CD₃OD) δ 28.41, 28.47, 28.80, 28.81, 58.7, 68.3, 84.19, 84.22, 95.8, 130.6, 145.7, 152.2, 152.7; HRMS (ESI–) calculated for C₁₇H₂₉N₂O₇S [M – H][–] 405.1701, found 405.1721 (error 4.9 p.p.m.).

Ser-AVS. To a solution of *N*⁶, *N*⁶-bis(*tert*-butoxycarbonyl)-2',3'-*O*-isopropylideneadenosine (4) (73 mg, 0.14 mmol, 1.1 equiv), vinylsulfonamide (3) (52 mg, 0.13 mmol, 1.0 equiv) and PPh₃ (56 mg, 0.21 mmol, 1.7 equiv) in THF (1 ml) at 0 °C, was added a solution of DIAD (42 μ l, 0.21 mmol, 1.7 equiv) in THF (1 ml) over 1 h using a syringe pump. The solution was gradually warmed up to 23 °C and stirred overnight. The mixture was filtered over a short pad of silica gel, which was washed with 20% EtOAc–hexanes (100 ml). The filtrate was concentrated to afford crude 5 (*R*_f = 0.45, 50:50 EtOAc–hexanes), which was used in the next step without further purification. To a solution of crude 5 from the previous step was added 80% aqueous trifluoroacetic acid (1 ml) at 0 °C. The solution was stirred for 6 h at 0 °C then concentrated. Recrystallization from 1:20 MeOH–Et₂O (5 ml) afforded the title compound (32 mg, 47%) as colourless film: [α] –10.3 (c 0.600, MeOH); ¹H NMR (600 MHz, CD₃OD) δ 3.30–3.39 (m, 2H), 3.67–3.70 (m, 1H), 3.83 (dd, *J* = 11.6, 4.1 Hz, 1H), 4.05–4.08 (m, 1H), 4.22–4.25 (m, 1H), 4.34–4.35 (m, 1H), 4.77–4.81 (m, 1H), 5.94 (d, *J* = 6.2 Hz, 1H), 6.70 (dd, *J* = 15.4, 6.5 Hz, 1H), 6.77 (d, *J* = 15.4 Hz, 1H), 8.27 (s, 1H), 8.29 (s, 1H); ¹³C NMR (150 MHz, CD₃OD) δ 45.8, 54.1, 62.3, 72.9, 74.8, 85.8, 91.7, 121.3, 134.8, 137.0, 143.2, 149.9, 151.3, 156.1; HRMS (ESI+) calculated for C₁₄H₂₂N₇O₆S [M + H]⁺ 416.1347, found 416.1339 (error 1.9 p.p.m.).

Kinetic analysis of AB3403. Substrate preference for the adenylation domain of holo-AB3403 was established by the pyrophosphate exchange assay⁴² allowing radiolabelled PP_i to be incorporated into ATP in the reverse reaction. One micromolar holo-AB3403 was added to 2 mM ATP, 0.2 mM NaPP_i, 50 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM MgCl₂, 0.15 μ Ci [³²P]PP_i, and 5 mM substrate. Reactions (100 μ l) were incubated for 10 min at 37 °C, then quenched with 0.5 ml 1.2% charcoal, 0.1 M unlabelled PP_i, and 0.35 M perchloric acid. The charcoal was pelleted by centrifugation, washed twice with 1 ml H₂O, and resuspended in 0.5 ml H₂O for scintillation counting.

To determine the apparent kinetic constants for ATP and glycine for the holo-AB3403 adenylation domain, the NADH⁺ consumption assay monitored at A_{340 nm} (refs 43, 44) with full-length AB3403. Hydroxylamine was used as a surrogate for the pantetheine in the second partial reaction to displace AMP for use in the coupled consumption assay⁴⁵. Standard reactions contained 50 mM HEPES (pH 7.5), 15 mM MgCl₂, 2 mM ATP, 3 mM phosphoenolpyruvate, 0.2 mM NADH⁺, 5 U myokinase, 5 U pyruvate kinase, 6.5 U lactate dehydrogenase, and 150 mM buffered hydroxylamine. Apparent kinetic constants were determined by varying concentrations of ATP or glycine with the one or the other in excess. Reactions were initiated by the addition of 0.001 mM enzyme. Calculations were done using PRISM software.

26. Adams, M. D. *et al.* Comparative genome sequence analysis of multidrug-resistant *Acinetobacter baumannii*. *J. Bacteriol.* **190**, 8053–8064 (2008).
27. Kapust, R. B. *et al.* Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **14**, 993–1000 (2001).
28. Doublé, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530 (1997).
29. Batty, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D* **67**, 271–281 (2011).
30. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
31. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protocols* **3**, 1171–1179 (2008).
32. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
33. Urzhumtsev, A., Afonine, P. V. & Adams, P. D. TLS from fundamentals to practice. *Crystallogr. Rev.* **19**, 230–270 (2013).
34. Roche, E. D. & Walsh, C. T. Dissection of the EntF condensation domain boundary and active site residues in nonribosomal peptide synthesis. *Biochemistry* **42**, 1334–1344 (2003).
35. Luft, J. R. *et al.* A deliberate approach to screening for initial crystallization conditions of biological macromolecules. *J. Struct. Biol.* **142**, 170–179 (2003).
36. Sundlov, J. A. & Gulick, A. M. Structure determination of the functional domain interaction of a chimeric nonribosomal peptide synthetase from a challenging crystal with noncrystallographic translational symmetry. *Acta Crystallogr. D* **69**, 1482–1492 (2013).
37. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
38. Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative staining and image classification – powerful tools in modern electron microscopy. *Biol. Proced. Online* **6**, 23–34 (2004).

39. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
40. Yang, Z., Fang, J., Chittuluru, J., Asturias, F. J. & Penczek, P. A. Iterative stable alignment and clustering of 2D transmission electron microscope images. *Structure* **20**, 237–247 (2012).
41. Ikeuchi, H., Meyer, M. E., Ding, Y., Hiratake, J. & Richards, N. G. A critical electrostatic interaction mediates inhibitor recognition by human asparagine synthetase. *Bioorg. Med. Chem.* **17**, 6641–6650 (2009).
42. Rusnak, F., Faraci, W. S. & Walsh, C. T. Subcloning, expression, and purification of the enterobactin biosynthetic enzyme 2,3-dihydroxybenzoate-AMP ligase: demonstration of enzyme-bound (2,3-dihydroxybenzoyl)adenylate product. *Biochemistry* **28**, 6827–6835 (1989).
43. Horswill, A. R. & Escalante-Semerena, J. C. Characterization of the propionyl-CoA synthetase (PrpE) enzyme of *Salmonella enterica*: residue Lys592 is required for propionyl-AMP synthesis. *Biochemistry* **41**, 2379–2387 (2002).
44. Reger, A. S., Carney, J. M. & Gulick, A. M. Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase. *Biochemistry* **46**, 6536–6546 (2007).
45. Wilson, D. J. & Aldrich, C. C. A continuous kinetic assay for adenylation enzyme activity and inhibition. *Anal. Biochem.* **404**, 56–63 (2010).
46. Reuter, D. C., McIntosh, J. E., Guinn, A. C. & Madera, A. M. Synthesis of vinyl sulfonamides using the Horner reaction. *Synthesis* **2003**, 2321–2324 (2003).

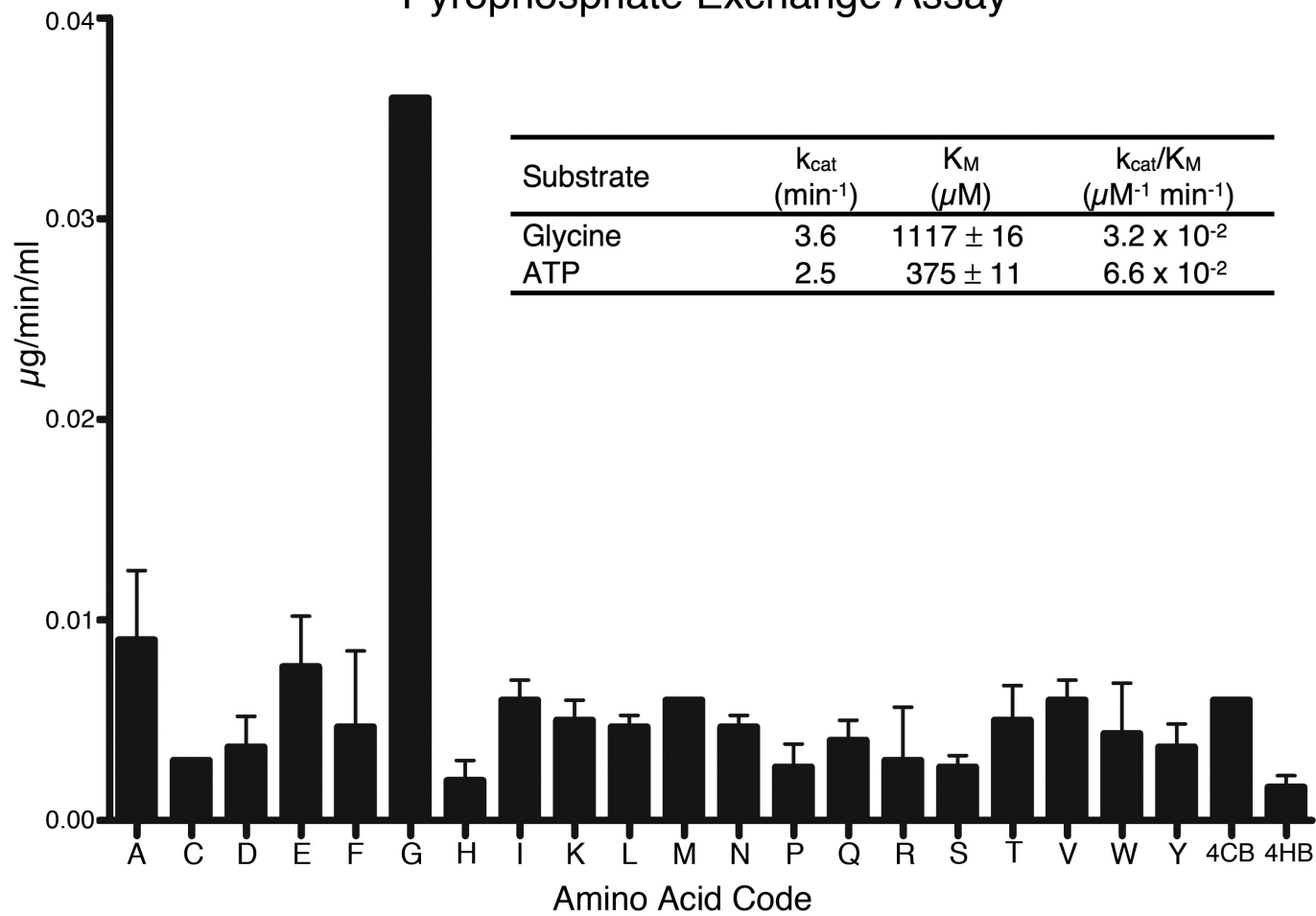
EntF	-----MSQHLPLVA-AQPGIWMAEKLSLPSAWVAHYVELTGEVDSP--LLARAVVAGLAQADTLRMRFTE--DNGEVWQWDDALTFELPEIDLRTNIDPHGTAQAL	100
AB3403	MNMLARLEPEVLSRHAISSEQLGIWYIQRLEPTCSAYNMVAVFVKVNSLGNKPIEILEAVMHYDPLLRVSMPA--NDQGEQLIWD--RVYPNIIIFSDA--RHIEASDLTQLV	109
SrfA-C	--MSQFSKQVQDMYLSLPMQEGMLFHAILNPGQSFYLEQITMKVKGSLNIK--CLEESNMVIMDRYDVFTVFIHEKVKRPVQVVLK--KRQFHIEEIDLTHLTGSEQTAKIN	108
EntF	MQAD-LQQDLRVDSGKPLVFHQLIQVADNRWYQRYHHLLVDGFSFPAITRQIANIYCTWLARGEPTASPFPFPADVVEEYQYQYRESEAWQRDAFVAEQRRQLPPASLS	211
AB3403	EQDT---KQPFDLTOPPLWRTHYECGQNHYYIAFVIHHAALDFWISGLLLRDVSKRGLVA---ESDAVNGIEFVQYADKQSSVDDDT--DESLIFWKNAKHAPHVHSIP	214
SrfA-C	EYKEQDKIRGFDLTDIPMRAAIFKKAESFEWVWSYHHIILDGWCFCGIVQDQLFKVYNALREQKPYSLPPVKPKDYIKWLEKQDK---QASLRWREYLEGEGFQTTFA	216
EntF	PAPLPGRSASADILRLKLEFDTGEFRQLATQLSGV--QRTD-LALALALWLGRCLNMDYAAGFIFMRR--LGSALTATGPVLNVLPLGIHIAAQETLPELATRLAAQLKK	319
AB3403	LDYPRPAVQKHGSSLVFRVSESVSSGLVNLAKDYEITLFGVLSGFYVLHLKSNENNLVIATPVAGR--LERSLRNALGQFVNTIAIHMDIDADQTLRQFTQQVQEQLRQ	324
SrfA-C	EQRKK--QKDGYPEKELLFSPSEAEKAFTELAKSQHTTLSTALQAVWSVLISRYQSSGDLAFGTVVSGRPAEIKGVEHVMGLFINVPRVRVLSGITFNGLLKRLQEQLSQ	327
EntF	MRRHQRYDAEQIVRDSGRA---AGDEPLFGPVLNIKVFYD-----QLDIPDVQAQHTLATG--PVNDLELALFPDVHGDLSIEILANKQRYDEPTLIQHAERL	413
AB3403	SLKHQKIAFSRVVEAVSPKRDGSINPLAQIGMFWERLGGMDEFKELLPIQTPTATLVGQDLTSGFVPRQEQGLDITLEMGGYQGLVGVLYKNTDLFSAQSAENMQVLL	436
SrfA-C	SEPHQYVPLDIQSQADQ-----PKLIDHIIVFENYPLQDA-----KNEESSENGDMVDVHVEK--SNYDLNLMASP--GDEMLIKLAYNENVFDEAFILRLKSQ	421
EntF	KMLIAQFAADPALLCGDVIDMLP-GEYAHVAQLNATQVEIPETTL SALVAEQAAKTPDAPALADARYLFSYREMEQVVALANLLRERGKPGDSVAVALPRSVFLTLALHA	524
AB3403	QAVLSMVAHPERKIVELDIADPYKDGQIFEARLKGKATDYAHDLFAMILKQIDERGDNHALTSDNDTVSYRELGHQIAGIAEYLAHAGITQDGRVGLMDRTALLPAAILG	548
SrfA-C	LTAIQQLIQNPDPQVSTLNLVDDREFFLLTGLNPPAQAHETKPLTYWFKAVNANPDAPALTYSGQTL SYRELDEEANRIARRLQKHGAGKGSVVALYTKRSLELVIGILG	533
EntF	IVEAGAAWPLDTPGDDRLKMLEDARPSLLITDDQLPRFSDV--PNLTSLCYNAPLT--PQGSAPLQLSQPHHTAYIIFTSGSTGRPKGMVQGTATVNRLLWMQNHYP	633
AB3403	IWAAGAAYVPLDPNPFPTERLQNIIEADAEKVLITQTELMQGLN---VSVPRLDINQAGV--VALEQVRETLAFGDIAAYMYTSGSTGKPKGVRIGHPISINFLSMNDRLQV	655
SrfA-C	VLKAGAAVLPVDPKLPEDRISYMLADSAACCLLTHQEMKEQAALPYTGTTLFIDDDQTRFEEQASDPATAIDPNDPAYIMYTSGTTGKPKGNITTHANIQLVKHVD--YMAF	644
EntF	TGEDVVAQKTPCSFDVSVWEFFWPFIAAGAKLVMAEPAHRDPLAMQFFAEYGVTTTHFVPSMLAAFVASLTPQTARQSCATLKQVFCSGEALPADLCREWQQLTGA--PLHN	744
AB3403	TTETQLLAITTYAFDISILELLIPLMYGGVVHVCPREVSQDGIQVLDYLNKASINVLQATPATWKMLLDSEWSG-----NAGLTALCGGEALDTILAEKLKGVG--CLWN	759
SrfA-C	SDQDTLVSNSYAFDAFTDFDYASMLNAARLIIADEHTLLDTERLTDILQENVMVFATTALFNLLTDAGEDWM-----KGLRCILFGGERASVPHVRKALRIMGPGKLIN	751
EntF	LYGPTEAADVSWYPAFGEELAQVRGSSVPIGYPVWNTGLRILDAMMHPVPPGVAGDLYLTGIQLAQGYLGRPDLTASRFIADPFAPGE--RMYRTGDVARWLDNGAVEYLG	855
AB3403	VYGPTEITTVWSAARTDA-----KYIDLGEPLANTQLYVLDQQRLLVPPGVMGELWAGGDLAVDYWRPELTDAQFRTLPSLPNAGRLYRTGDKVLRDGLTHHGR	864
SrfA-C	CYGPTEGTVFATAHVHDLPSDI---SSLPIGKPISNASYILNEQSQQLPFGAVGELCISGMGVSCKGYNRADLTKEKFIENPFKPGE--TLYRTGDLARWLDPGTIEYAGR	859
EntF	SDDQLKIRGQRIELGEIDRVMAQLPDVEQAVTHACVINQAATGGDARQLVGLVVSQSLPLDTSALQAQLRETLPPHMPVVLQLPQLPLSANGKL--DRKALPLPELK-A	965
AB3403	LDFQVKIRGFRIELGEIENVLKQIDGITDAVVLVKT-----GDNDQKLVAAYTG---QELDIAGLKKNLQIHLPAYMVPFAIRLDFEPMTANKKL--DRKAFPEPIFE-Q	965
SrfA-C	IDQVKIRGHRIELEIEEKQLQEYPGVKDAVVADRHE-----SGDASINAYLVN--RTQLSAEDVKAHLKKQLPAYMVPQTFTFLDELPLTTNGKVNKR--LLPKPDQDQL	962
EntF	QAPGRAPKAGSETIIAAAFSSLLGCDVQDADADFFALGGHSLLAMKLAQLSRQVARQVTPGQVMVASTVAKLATIIDAEEDSTRMGFETILPLREGNG--PTLFCFHPAS	1075
AB3403	SNDYVAPRDPPIELCTTFEQILSVKRVGIHDDFFELGGHSLLAVKLVNHLKKAFTGELSVALLAQYSTVERLGEIIRENKE----IKPSIVIELRR--GYEQPLWLPHPIG	1072
SrfA-C	AEEWIGRNMEEETIAQIWSEVLGRKQIGIHDDFFALGGHSLKAMTAASRKELGIDLVPKLLFEAPTIAGISAYLKNKGS---DGLQDVTIMNQDQE--QIIFAFPPL	1068
EntF	GFAWQFSVLRYLDPOWSIIIGIQRPNPGMQTAANLDEVCEAHLATLLEQQHPGYPYLLGYSLGGLAQGIAARLRARGEQVAFGLGLDTPWPETQNWQEKEANGLDPEVL	1187
AB3403	GSTFCYMLSRHLNPNRLRAIQSPGLIEADAEEVAIEEMATLYIAEMQKMQPGPYFLGGWCFGGAIAYEISRQLRQMGQVTVGIVMIDTRAPIENVPEDADDAMLLSWF	1184
SrfA-C	GYGLMYQLNSSL--PSYKLCADFIEE-----EDRLDRYADLIQKLQPEGLTLFGYSAGCSLAFEAALKLEEGRIVQRIIMVDSYKKGVSDDLGRTVESDVEAL	1169
EntF	AEINR-----EREFLAAQGGSTSTELFTTIEGNYADAVRLTTAHSV--PFDGKATLFVAERTLQEGMSPERAWSPWIAELDIYRQDCAH	1272
AB3403	ARDLAAPYGGKLTIPAAQYRLSPDQMFHVLKEAKAINVPLDADPSDFRLYFDTYLANGIALQTYFPE--PEDFPILLVAKDEQEDFGESLGDVQLVKDTLTQVDLPDGH	1295
SrfA-C	-----EALMNVNRDNEALNSEAVKHGLKQKTHAFYSYYNLISTGQVKADIDLLTSGADFIPPELWASWEEATTGVYRMKRGFGTH	1247
EntF	VDIISPGTFEKIGIPIRATLNR*	1293
AB3403	SSIMYAENVVGLADIDQMYPIA*	1317
SrfA-C	AEMLQGETLDRNAELLEFLNTQTVTVS*	1275

	EntF	AB3403	SrfA-C
Condensation	1-429	1-452	1-437
Adenylation	446-964	467-964	452-962
PCP	972-1052	972-1048	969-1044
Thioesterase	1055-1293	1051-1317	1047-1275
Hinge	Asp857	Asp866	Asp861
PPant	Ser1006	Ser1006	Ser1003

Extended Data Figure 1 | Structure-based alignment of EntF, AB3403, and SrfA-C. Condensation, adenylation, PCP, and thioesterase domains are represented with bars in grey, pink, green-cyan, and blue. Conserved motifs and catalytically important residues are highlighted with the same colours, including the HHxxxD motif of the condensation domains, the aspartic acid hinge that separates the N- and C-terminal subdomains of the adenylation domain, the GGHS motif that is the site of pantetheinylation

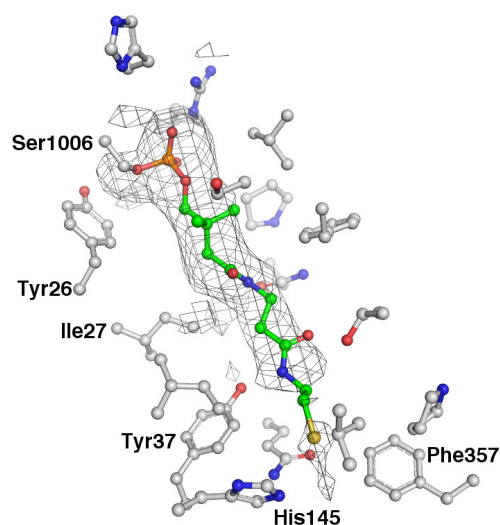
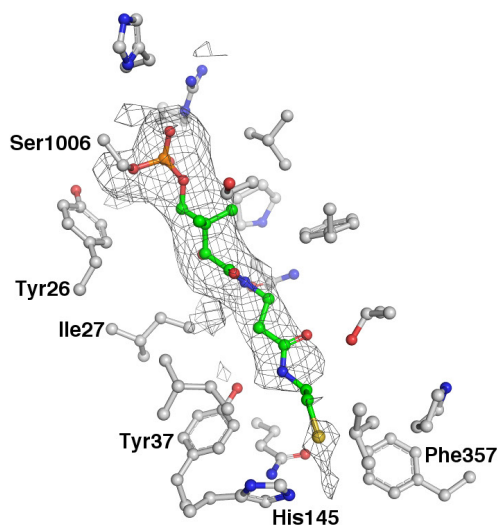
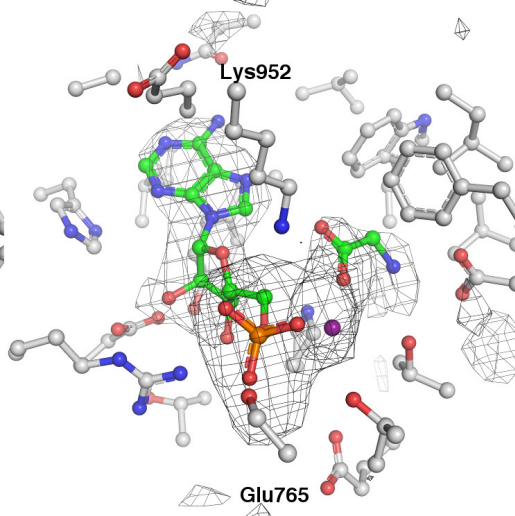
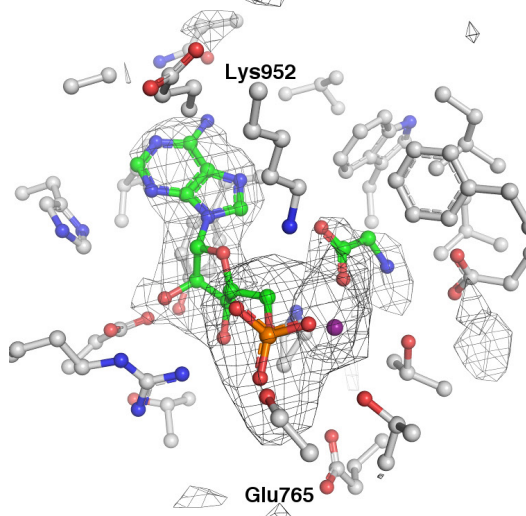
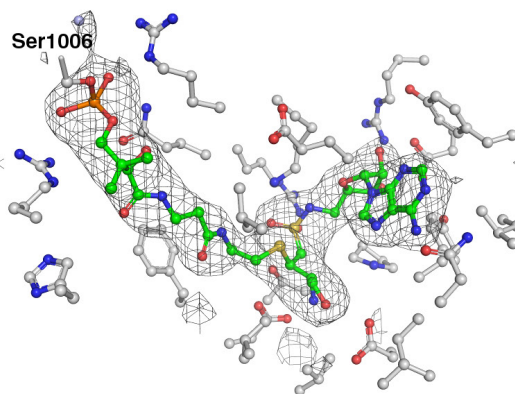
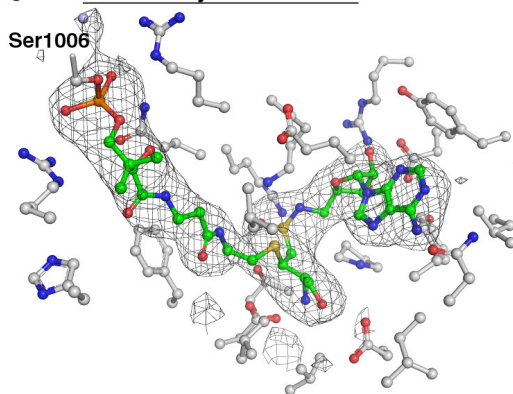
in the PCP, and the catalytic nucleophile of the thioesterase domain. The SrfA-C, AB3403, and EntF proteins share approximately 26% sequence identity. The adenylation and PCP domains are more well-conserved, sharing ~35% identity, whereas the condensation (21%) and thioesterase (25%) domains are less well conserved. Domain boundaries are described in the table below.

Pyrophosphate Exchange Assay



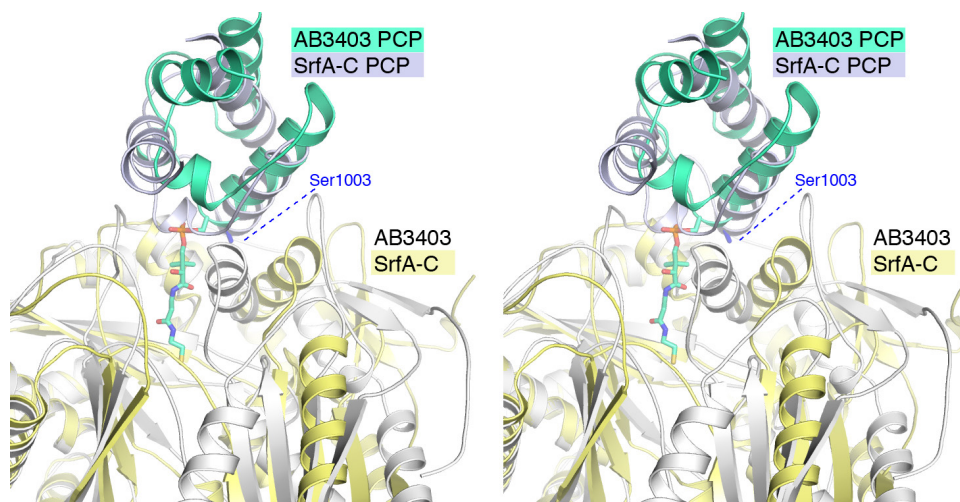
Extended Data Figure 2 | Substrate specificity of full-length AB3403. Amino-acid specificity of AB3403 was recorded for all 20 proteinogenic amino acids, as well as 4-chlorobenzoate (4CB) and 4-hydroxybenzoate (4HB). Average values and standard deviations are shown for three

replicates with each substrate; results were recorded as micromoles of radiolabelled ATP incorporated per minute per milligram of enzyme. Apparent kinetic constants are also shown for ATP and glycine calculated from duplicate measurements for four to six substrate concentrations.

a AB3403 Condensation Domain**b** AB3403 Adenylation Domain**c** EntF Adenylation Domain

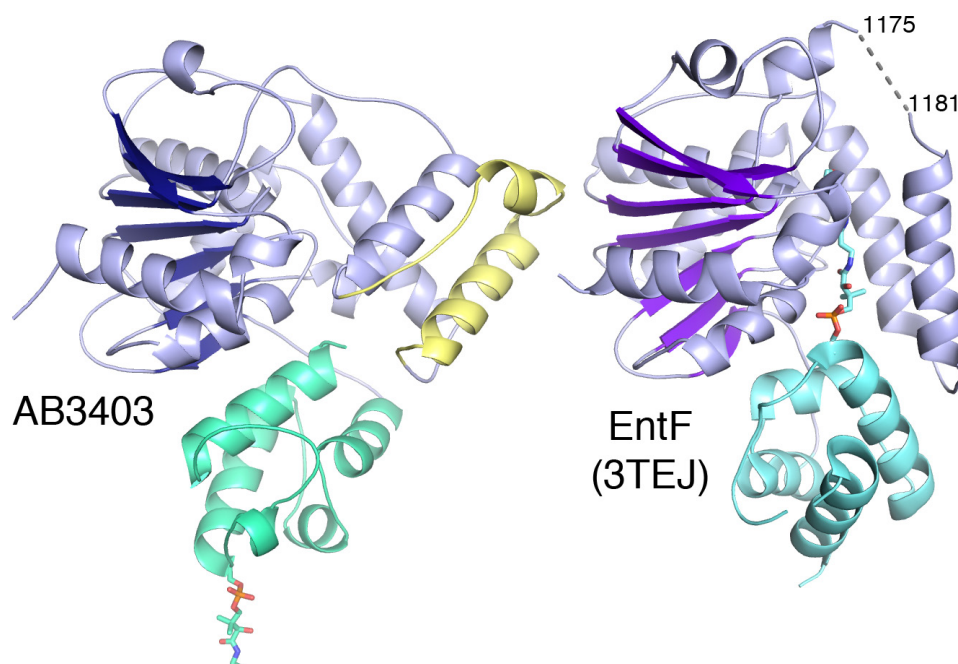
Extended Data Figure 3 | Stereo representations of electron density figures shown in Fig. 2. To better visualize the active sites and electron density quality, stereo figures are included in the extended data. In all panels, density is shown with coefficients of the form ($F_o - F_c$) calculated before inclusion of ligands and contoured at 3σ . **a**, Stereo representation of electron density of AB3403 condensation domain shows the phosphopantetheine on Ser1006 approaching His145 within the condensation domain pocket. Inhibitor carbon atoms in green, carbons of residues within 5 Å of inhibitor in grey, nitrogen in blue, oxygen in red, sulphur in yellow, and water in light blue. **b**, Electron density of the

nucleotide binding pocket of AB3403 bound to glycine and AMP. Stereo representation of electron density shows the AMP, glycine, and Mg^{2+} present in the active site of the adenylation domain. Ligand carbon atoms are in green, carbons of residues within 5 Å of inhibitor in grey, nitrogen in blue, oxygen in red, phosphorus in orange, and the Mg^{2+} cofactor in purple. **c**, Stereo representation of the electron density shows the phosphopantetheine on Ser1006 covalently attached to the Ser-AVS inhibitor in the active site of the adenylation domain. Inhibitor carbon atoms in green, carbons of residues within 4 Å of inhibitor in grey, nitrogen in blue, oxygen in red, phosphorus in orange, sulphur in yellow, and water in light blue.



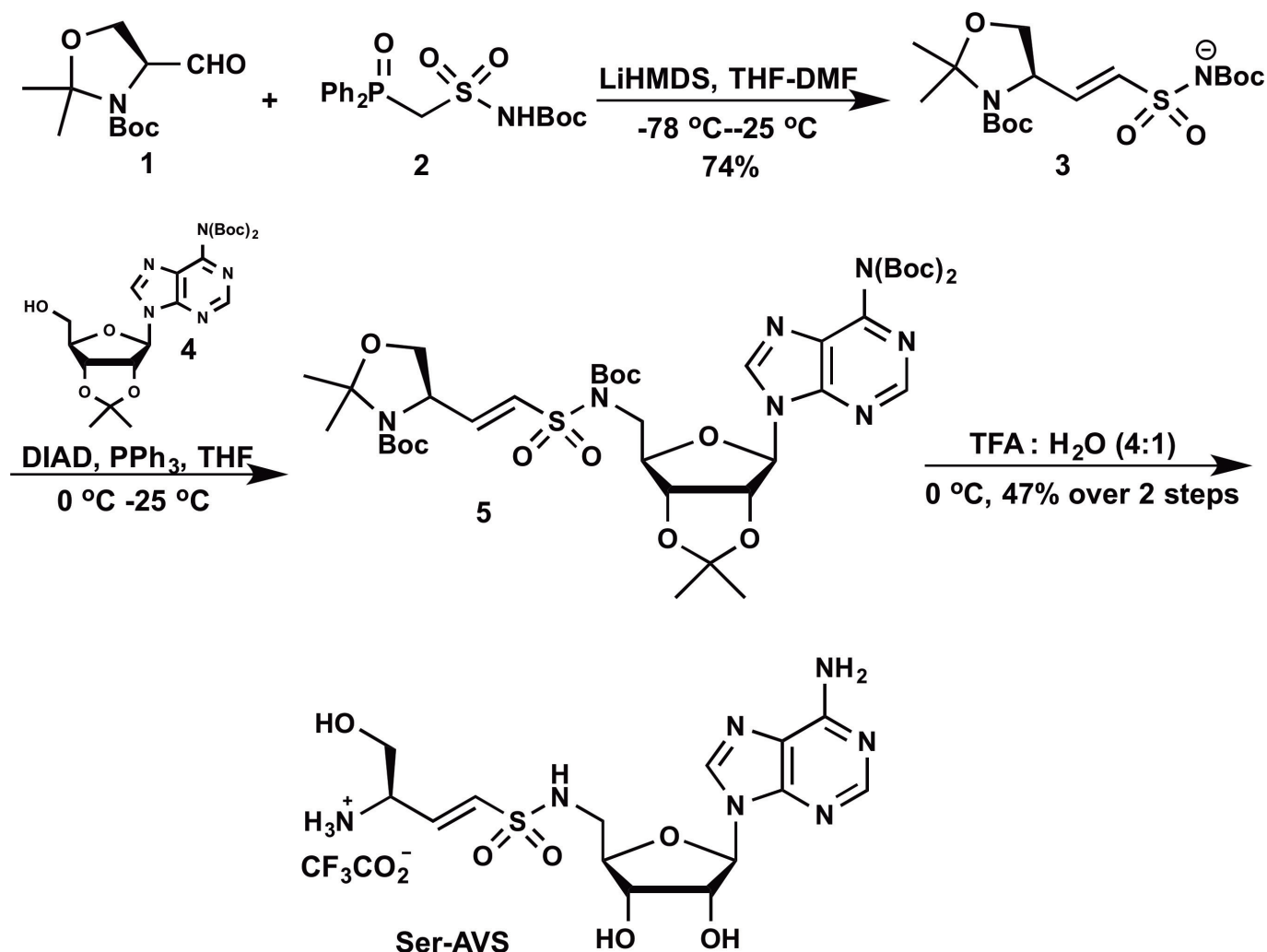
Extended Data Figure 4 | Comparison of AB3403 and SrfA-C PCP-condensation domain interaction. Stereo representation illustrating different orientations of the PCP domains of SrfA-C and AB3403 relative to the condensation domains with which they interact. AB3403 is shown

with a white condensation domain and a green-cyan PCP. SrfA-C is shown with a yellow condensation domain and a pale blue PCP. The pantetheine of AB3403 is shown bound to Ser1006. The position of Ser1003, mutated to an alanine residue in SrfA-C, is also highlighted.



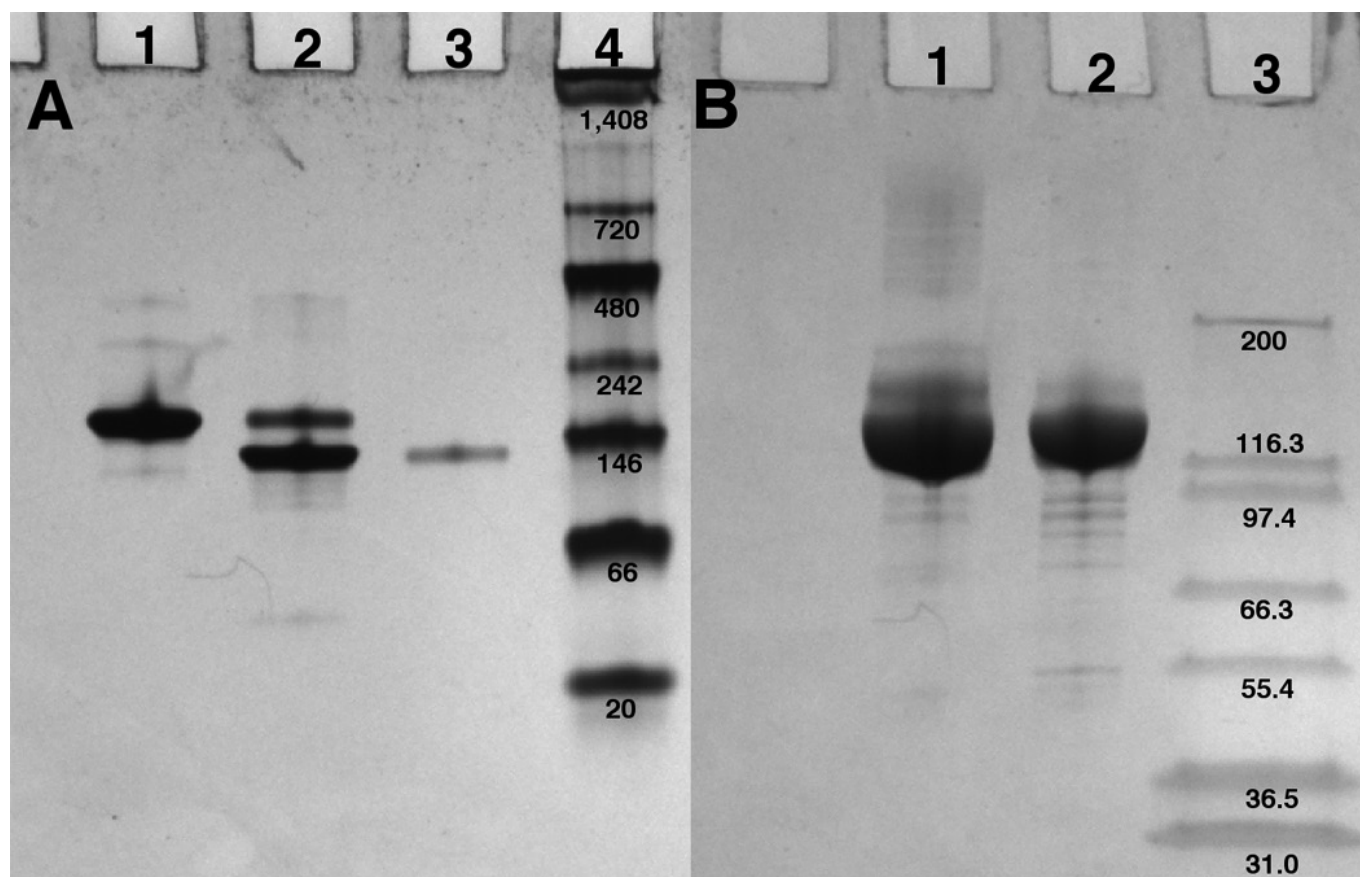
Extended Data Figure 5 | Comparison of AB3403 thioesterase domain to the functional PCP–thioesterase interaction. Stereo representation of the thioesterase (blue) domain of AB3403 interacts with the back face of the PCP domain in AB3403. The functional interaction between the

EntF thioesterase domain and its holo-PCP, trapped crystallographically, illustrates that the same face of the thioesterase domain interacts functionally (PDB 3TEJ). A 28-residue insertion of AB3403 is coloured yellow.



Extended Data Figure 6 | Synthesis of Ser-AVS. The Ser-AVS probe was synthesized following similar protocols described elsewhere^{41,46}. Garner's aldehyde **1** was coupled with **2** using LiHMDS to exclusively furnish the

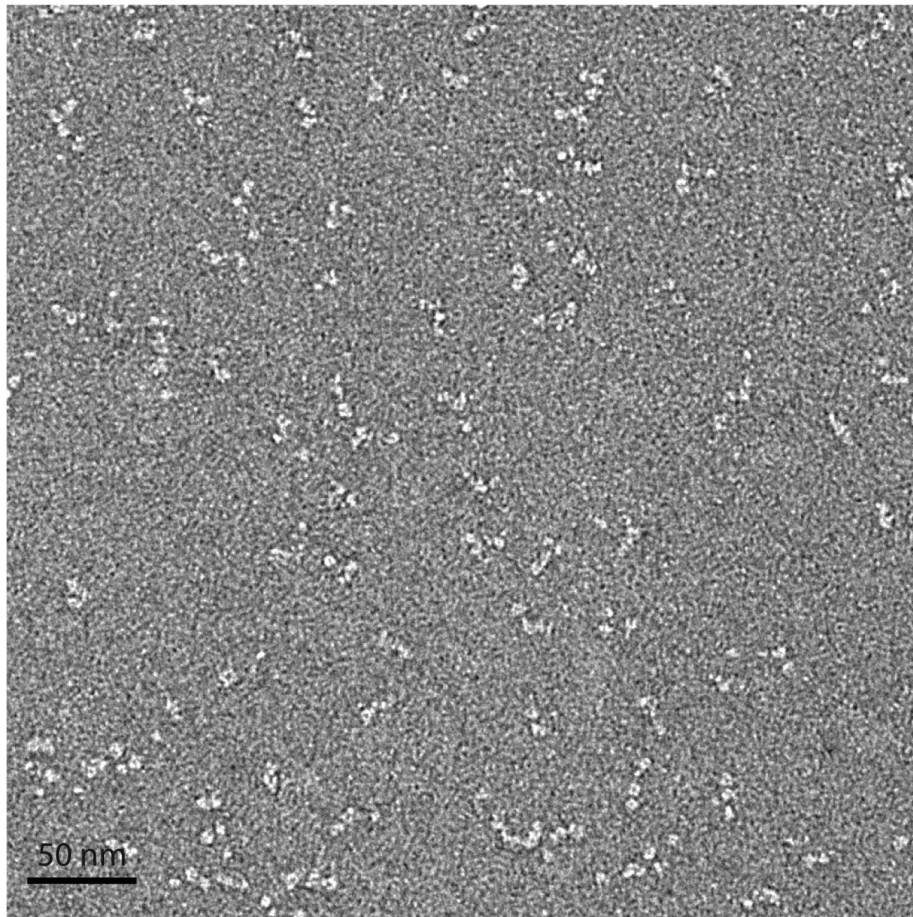
(*E*)-vinylsulfonamide **3**. Mitsunobu coupling of **3** with bis-Boc adenosine **4** afforded **5**, which was globally deprotected using 80% aqueous trifluoroacetic acid to yield Ser-AVS.



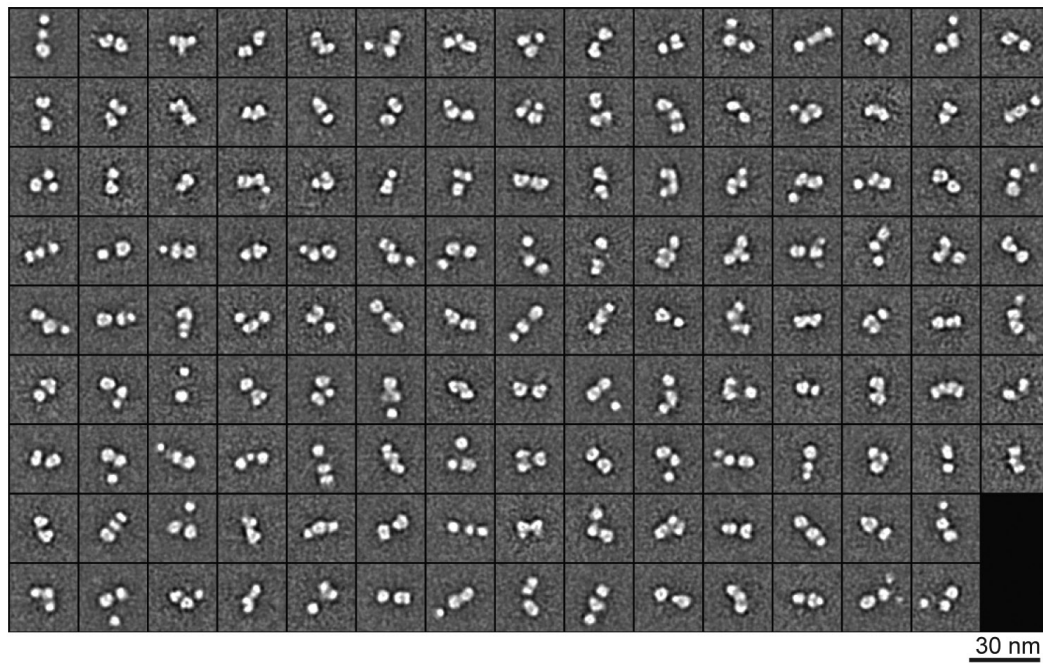
Extended Data Figure 7 | Electrophoretic mobility of EntF. a, Native gel electrophoresis. Lane 1: EntF. Lane 2: EntF incubated with fourfold molar excess of Ser-AVS inhibitor. Lane 3: EntF Crystals. Lane 4: novex NativeMark labelled in kilodaltons. **b,** Denaturing gel electrophoresis using loading buffer with SDS and β -mercaptoethanol. Gel lane 1: EntF.

Lane 2: EntF incubated four times with Ser-AVS inhibitor. Lane 3: Life Technologies Mark12 labelled in kilodaltons. The native gel shows the inhibited EntF in a more compact conformation compared with EntF without the inhibitor.

a



b



Extended Data Figure 8 | Negative-stain electron microscopy analysis of EntF. **a**, Raw electron microscopy image of negative-stained EntF. **b**, Class averages of EntF particles.

Extended Data Table 1 | Diffraction data statistics and refinement statistics for AB3403

	SeMet_AB3403	AB3403	Liganded AB3403
PDB Code		4ZXH	4ZHI
Beamline	APS 23-ID-B	APS 23-ID-B	APS 23-ID-B
Wavelength (Å)	0.9793	0.9796	1.0332
Space group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Unit cell a, b, c (Å)	116.19 116.19 348.61	116.19 116.19 348.61	116.10 116.10 342.02
Molecules per ASU	1	1	1
Resolution range (Å)	29.75–3.35 (3.52–3.35)	49.80–2.70 (2.79 – 2.70)	45.03–2.90 (3.00 – 2.90)
Total reflections	137397 (16096)	416923 (21743)	257325 (25582)
Unique reflections	34599(4299)	66559 (6495)	52900 (5187)
Multiplicity	4.0 (3.7)	6.3 (3.4)	4.9 (4.9)
Completeness (%)	98.9 (94.6)	99.96 (100.00)	99.99 (100.00)
Mean I/sigma(I)	11.9 (3.5)	9.91 (2.49)	8.47 (2.19)
R _{merge}	0.090 (0.359)	0.125 (0.511)	0.130 (0.641)
R _{meas}	0.116	0.143	0.162
CC1/2	0.993 (0.798)	0.991 (0.685)	0.991 (0.635)
CC*	0.998 (0.937)	0.998 (0.902)	0.998 (0.881)
Structure Refinement			
R _{factor}		0.179 (0.248)	0.174 (0.307)
R _{free}		0.234 (0.322)	0.225 (0.369)
No. atoms		10301	10198
RMSD bond distances (Å)		0.009	0.009
RMSD bond angles		1.18	1.18
Wilson B-factor (Å ²)		41.15	50.83
Average B-Factor (Å ²)			
Protein		46.00	53.60
Ligand		54.40	55.30
Ramachandran analysis			
Favored (%)		97.0	96.0
Allowed (%)		2.3	3.4
Outliers (%)		0.7	0.6
Molprobability ClashScore		4.51	6.75

Values in parentheses are for the highest resolution shell.

Extended Data Table 2 | Diffraction data statistics and refinement statistics for EntF

	SeMet EntF	EntF
PDB Code		4ZXJ
Beamline	APS 23-ID-B	APS 23-ID-B
Wavelength	0.9796	1.0332
Space group	P4 ₁ 2 ₁ 2	P4 ₁ 2 ₁ 2
Unit cell a, b, c (Å)	127.55 127.55 186.72	127.71 127.71 186.94
Molecules per ASU	1	1
Resolution range (Å)	60 – 2.9 (3.0 – 2.9)	81.31 – 2.8(2.9 – 2.8)
Total reflections	152578 (15129)	175399 (17288)
Unique reflections	34693 (3380)	38753 (3800)
Multiplicity	4.4 (4.5)	4.5 (4.5)
Completeness (%)	99.66 (99.41)	99.96 (99.92)
Mean I/sigma(I)	9.86 (2.49)	9.85 (2.11)
R-merge	0.1153 (0.6165)	0.0979 (0.6484)
R-meas	0.1312	0.1109
CC1/2	0.995 (0.598)	0.997 (0.629)
CC*	0.999 (0.865)	0.999 (0.879)
Structure Refinement		
R _{factor}		0.183 (0.290)
R _{free}		0.230 (0.324)
No. protein/ligand atoms		7898/49
RMSD bond distances (Å)		0.008
RMSD bond angles		1.23
Wilson B-factor (Å ²)		62.88
Average B-Factor (Å ²)		
Protein		74.6
Ligand		49.8
Condensation N-terminal Subdomain		90.5
Condensation C-terminal Subdomain		84.8
Adenylation N-terminal Subdomain		54.9
Adenylation C-terminal Subdomain		54.8
PCP Domain		78.8
Ramachandran analysis		
Favored (%)		95.0
Allowed (%)		4.13
Outliers (%)		0.87
Molprobability ClashScore		6.03

Values in parentheses are for the highest resolution shell.

Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase

Janice M. Reimer^{1*}, Martin N. Aloise^{1*}, Paul M. Harrison² & T. Martin Schmeing¹

Nonribosomal peptide synthetases (NRPSs) are very large proteins that produce small peptide molecules with wide-ranging biological activities, including environmentally friendly chemicals and many widely used therapeutics¹. NRPSs are macromolecular machines, with modular assembly-line logic, a complex catalytic cycle, moving parts and many active sites^{2,3}. In addition to the core domains required to link the substrates, they often include specialized tailoring domains, which introduce chemical modifications and allow the product to access a large expanse of chemical space^{3,4}. It is still unknown how the NRPS tailoring domains are structurally accommodated into megaenzymes or how they have adapted to function in nonribosomal peptide synthesis. Here we present a series of crystal structures of the initiation module of an antibiotic-producing NRPS, linear gramicidin synthetase^{5,6}. This module includes the specialized tailoring formylation domain, and states are captured that represent every major step of the assembly-line synthesis in the initiation module. The transitions between conformations are large in scale, with both the peptidyl carrier protein domain and the adenylation subdomain undergoing huge movements to transport substrate between distal active sites. The structures highlight the great versatility of NRPSs, as small domains repurpose and recycle their limited interfaces to interact with their various binding partners. Understanding tailoring domains is important if NRPSs are to be utilized in the production of novel therapeutics.

Tailoring domains embedded within NRPSs are vital for the production and bioactivity of the nonribosomal peptide (NRP) products of these synthetases⁴. Tailoring domains exist in addition to the core NRPS adenylation (A), peptidyl carrier protein (PCP) and condensation (C) domains, which a module requires to add an amino acid to the growing NRP: the A domain selects, activates and transfers the substrate amino acid to the PCP domain, which transports it to the C domain for peptide bond formation¹ (Fig. 1 and Extended Data Fig. 1). Tailoring domains are common in NRPSs^{3,4}, for example, cyclosporin synthetase contains methyltransferase domains⁷; daptomycin (Cubicin) synthetase, epimerization domains⁸; bacitracin (BACiM) synthetase, a heterocyclization domain⁹; valinomycin synthetase, ketoreductase domains¹⁰; bleomycin synthetase, an oxidase domain¹¹; and soframycin synthetase, a reductase domain¹². These domains enable key functionalities of the NRP by, for example, providing protease resistance, enabling novel interactions, improving affinity by limiting NRP conformational flexibility, or allowing the NRP to assume its active conformation. Linear gramicidin synthetase (comprised of LgrA, LgrB, LgrC, and LgrD) was previously shown to contain an active formylation (F) domain as the first domain of its F–A–PCP initiation module^{5,6} (Fig. 1). F domains are homologous to formyltransferase proteins that modify substrates in three diverse pathways: ribosomal translation¹³, purine anabolism¹⁴ and bacterial outer membrane synthesis¹⁵. The LgrA initiation module must formylate its substrate for linear gramicidin synthesis

to proceed⁶ (Fig. 1), and this formyl group is essential for the clinically important antibacterial activity of gramicidin¹⁶. Gramicidin molecules form head-to-head dimers through the formyl group to make a β -helical pore in gram-positive bacterial membranes. This pore allows free passage of monovalent cations, destroying the ion gradient and killing bacteria.

We have determined four independent crystal structures of the initiation module of LgrA at 2.5, 2.6, 2.8, and 3.8 Å resolutions (Extended Data Table 1 and Extended Data Fig. 2), showing four different functional conformations: the A domain open (substrate binding), the A domain closed (adenylation), thiolation and formylation states (Fig. 2, Extended Data Fig. 3 and Supplementary Video 1). Our data augment the existing structural knowledge of NRPSs (reviewed in ref. 2), by visualizing the structure of an NRPS module that includes a tailoring domain, showing how the tailoring domain is incorporated into and used as part of an NRPS, observing several functional states (open, closed, thiolation) in a single protein rather than over different excised

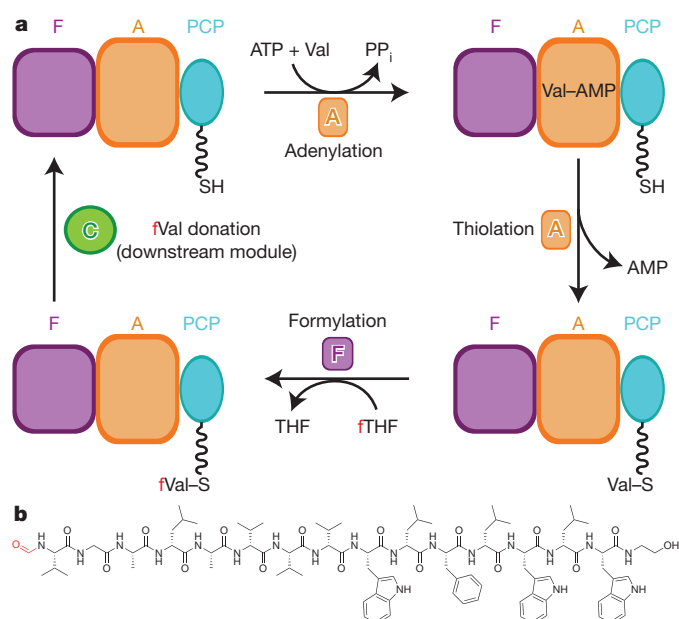


Figure 1 | A schematic of the action of the linear gramicidin synthetase initiation module. **a**, The F–A–PCP initiation module is the first module of LgrA, the dimodular F–A–PCP–C–A–PCP–E* NRPS protein in the LgrA–E synthetic cluster (E*, inactive epimerization domain). The initiation cycle begins with valine selection and adenylation followed by thiolation onto the PPE arm of the PCP domain. The F domain formylates PCP–PPE–Val before it is brought to be the donor in the condensation reaction of the downstream module. **b**, Chemical structure of linear gramicidin A.

¹Department of Biochemistry, McGill University, 3649 Promenade Sir-William-Osler, Montréal, Québec H3G 0B1, Canada. ²Department of Biology, McGill University, 1205 Dr Penfield Avenue, Montréal, Québec H3A 1B1, Canada.

*These authors contributed equally to this work.

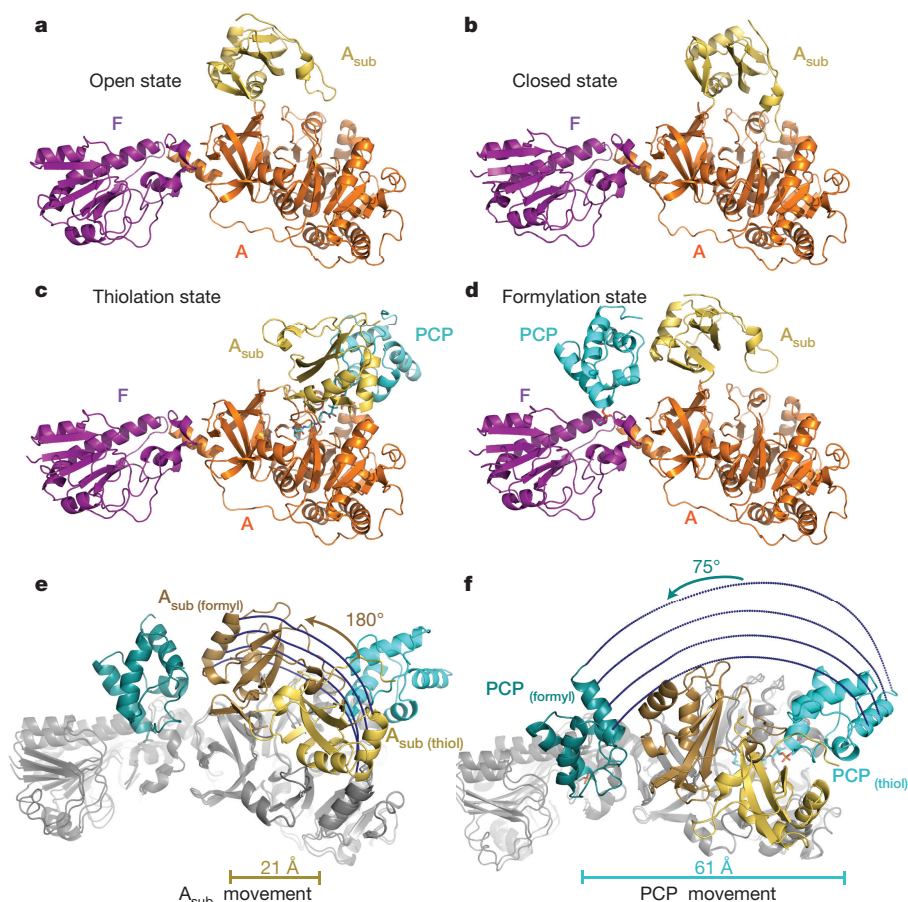


Figure 2 | Crystal structures representing the steps of the synthesis cycle in the LgrA initiation module. a–d, The F–A–PCP LgrA initiation module in open (a), closed (b), thiolation (c) and formylation (d) states. (The PCP domain is not necessary for the open and closed states and is disordered in b and c.) The transition between thiolation and formylation

states requires large rigid body movements of both the A_{sub} and PCP domains. e, The PCP domain rotates 75° and translocates its centre of mass by 61 Å. The PPE arm attachment point, Ser729, moves 52 Å, and some residues move >80 Å. f, The A_{sub} domain rotates 180° and translocates its centre of mass by 21 Å.

mono- and didomains^{2,17–21}, and visualizing a novel functional state (formylation).

The F domain is connected to the rest of the F–A–PCP LgrA initiation module through an interface with the A domain (Fig. 3) that buries 830 Å² of surface area. This is distinct from the C–A interface in C–A–PCP elongation and termination modules²² (Extended Data Fig. 4). The F–A interface appears sufficient to maintain these domains in a very elongated conformation (Fig. 2). Across all nonequivalent molecules in the crystals, the relative orientation between the two domains varies only by ~5°, and our small angle X-ray scattering analysis indicates that this extended conformation is representative of the initiation module in solution (Extended Data Fig. 5). This architecture means that the adenylation active site and the formylation active site are always ~50 Å apart, necessitating that the valine substrate travels a large distance between subsequent steps in synthesis. Accordingly, positions of the PCP domain and the A_{sub} domain (C-terminal portion of the A domain) change markedly in the progression of the module through functional states.

The NRPS assembly-line process (Fig. 1, Extended Data Fig. 1 and Supplementary Video 2) begins with ATP and valine binding to an open conformation of the A domain²⁰ (Fig. 2a and Supplementary Video 1). The A domain closes on substrate binding by rotating the A_{sub} by ~30° to catalyse formation of the valine adenylate^{17,18} (Fig. 2b). Next, the thiolation reaction transfers the valine from the adenylate to the thiol of the PCP domain phosphopantetheine arm (PPE). We accessed this state by attaching a non-hydrolysable analogue²³ of the product of the reaction, Val–NH–PPE, to the PCP domain. The resulting structure shows the known 140° rotation of

the A_{sub}^{19,21} and the product Val–NH–PCP still bound to the active site (Fig. 2c and Extended Data Fig. 2e). The PCP then transports its valine 50 Å between the A and F domain active sites to accept a formyl group. Our next structure (Fig. 2d) shows that to achieve this, the PCP domain makes a very large movement of a rigid ~75° rotation and 61 Å translocation (Fig. 2e). The ~10-residue linker between the A and PCP domains is not nearly sufficient to span the 55 Å travelled by the first residue of the PCP domain; accordingly, the A_{sub} domain undergoes a full 180° rotation and 21 Å translocation to allow the PCP domain to bind the F domain (Fig. 2f). There, Val–PCP accepts a formyl (f) group from the donor cofactor formyltetrahydrofolate (fTHF) (Extended Data Fig. 2f) onto its amino group^{5,6}. The PCP then moves the formyl–valine to the next module, where the condensation domain of that module will catalyse peptide bond formation between fVal–PCP and its Gly–PCP2, making the first peptide bond of linear gramicidin and liberating the PCP to participate in the next round of reactions⁶ (Supplementary Video 2).

How did the F domain become a functional NRPS domain? The F domain of LgrA was fused into an existing NRPS⁵, and we suggest that the pre-transfer source was a single-domain formyltransferase from a distantly related bacterium with a signature of missing helix α2 and strand β3. As the high incidence of horizontal transfer (Extended Data Fig. 6) is consistent with conferring a competitive advantage, and bacteria possessing formyltransferases similar to the F domain also have canonical transfer RNA and phosphoribosylglycinamide formyltransferases, it is likely that the pre-transfer formyltransferase performed the remaining known formyltransferase function, that is, sugar formylation for cell wall synthesis. After fusion, the F–A–PCP initiation

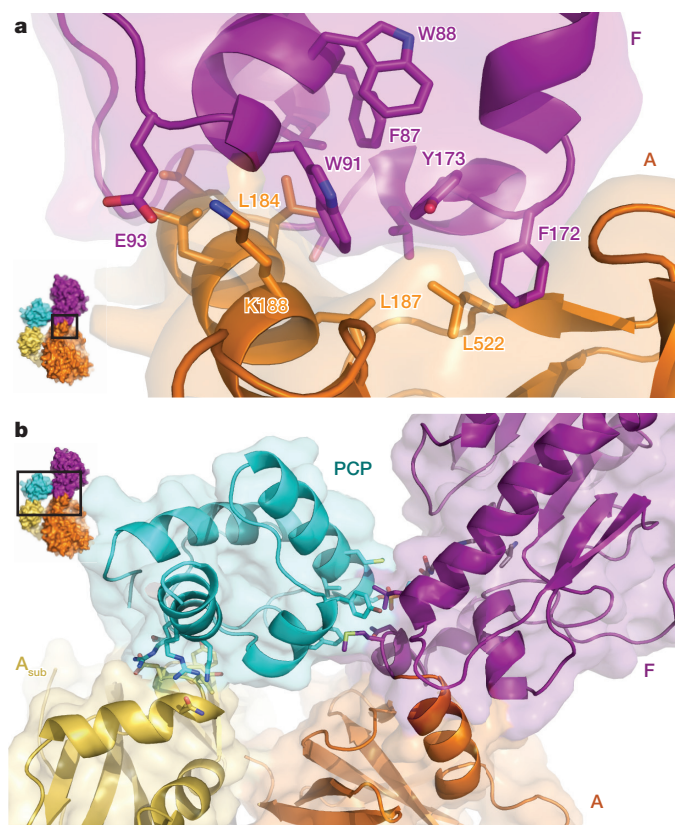


Figure 3 | Interdomain interfaces of the initiation module. **a**, The F domain is fused onto the A domain and forms a small hydrophobic core (Extended Data Fig. 8). **b**, Interaction of the PCP domain with A_{sub} and F domains in the formylation state. The A_{sub} domain creates an electrostatic platform for the PCP domain. The PCP domain binds to F domain hydrophobic residues Leu127 (often Lys or Glu in formyltransferases) and Met178 (in the C terminus of the F domain that is not similar to formyltransferases). The PPE phosphate interacts with Arg170 (often Glu, Ser or Asn in formyltransferases) and Asn177 (usually Glu, Asp or Met in formyltransferases).

module evolved rapidly (Extended Data Figs 6 and 7). The fold of the first 171 amino acids of LgrA is conserved with the sugar formyltransferases, leaving only residues 172–179, including a single α -helix, as a new structural element and link to the A domain (Extended Data Fig. 8). A ‘landing pad’ evolved to include a hydrophobic patch for binding the PCP domain, and positive residues and hydrogen bond donors to interact with the PPE phosphate (Fig. 3 and Extended Data Fig. 8). The F–PCP interaction places the PPE attachment point, Ser729, an ideal 16 Å away from the fTHF in the conserved formyltransferase active site. Notably, this positions the Val–PPE exactly in the thymidine diphosphate (dTDP)–sugar binding site of sugar formyltransferases^{5,15,24} (Fig. 4a). The similar length and hydrophilic nature of the dTDP–sugar and Val–PPE probably enabled the F domain to formylate Val–PCP soon after the fusion event, before formylation was absolutely required for downstream peptide synthesis to proceed.

The PCP domain interaction with the F domain is quite minimal, and accordingly, the A_{sub} domain donates an additional binding interaction in the formylation state⁵ (Fig. 3b). This is reminiscent of methionyl-tRNA^{Met} formyltransferase (FMT), the essential bacterial two-domain formyltransferase that uses its C-terminal domain (FMT_{CTD}) to present the methionyl-tRNA^{Met} to the formyltransferase active site¹³ (Fig. 4b, c). This functional convergent evolution presents another interesting parallel to ribosomal translation, the completely separate macromolecular system that also synthesizes peptides. In both LgrA and the ribosome, a mobile carrier macromolecule (PCP domain/tRNA) covalently (through thioester/ester bonds) transports an amino

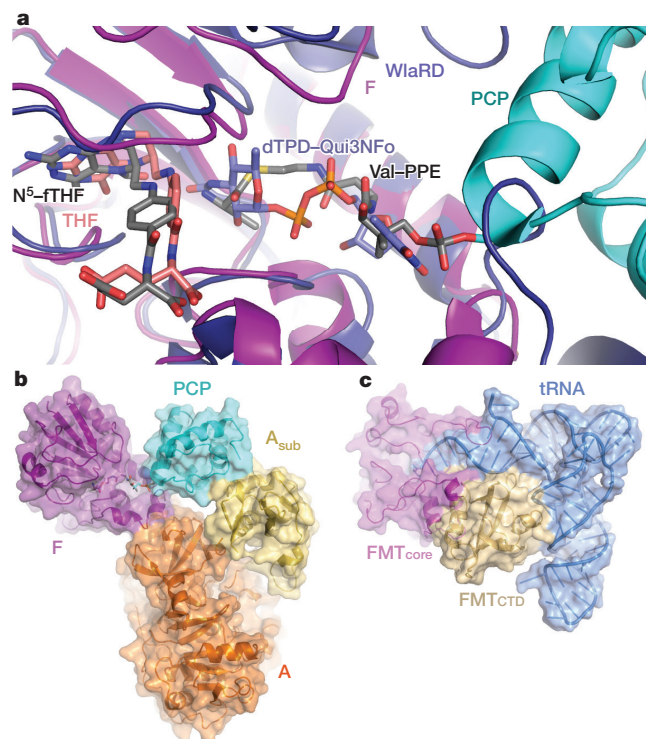


Figure 4 | Comparisons of the F domain to sugar and tRNA formyltransferases. **a**, The binding mode for the PPE arm to the F domain is similar to that of sugar–dTDP in sugar formyltransferases (protein WlaRD, PDB 4LY3 (ref. 15)). Note that the valine and most of the PPE arm (carbons shown in grey) are modelled, as they are not visible in electron density maps at 3.8 Å resolution. Qui3Nfo, 3,6-dideoxy-3-formamido-D-glucose. **b**, The A_{sub} domain emulates the positioning role of the FMT_{CTD} in methionyl-tRNA^{Met} formyltransferases (PDB 2FMT (ref. 13)). Excluding the PPE arm, the PCP domain buries only 279 Å² of F domain surface. The A_{sub} provides an additional 345 Å² of interaction surface to position the PCP domain.

acid to a formyltransferase enzyme (F domain/FMT), where the carrier is oriented by a positioning domain (A_{sub} /FMT_{CTD}) to allow formylation before acting as the first donor substrate for a peptidyl transferase enzyme (C domain/large ribosomal subunit).

Observing the same protein in these conformations, including the novel formylation conformation, highlights the versatility of the small domains in NRPSs^{2,3,18,23}. The small ~100-residue A_{sub} has three distinct roles in the cycle: providing catalytic residues for the adenylation reaction¹⁸; positioning the PCP for the thiolation reaction and later for the formylation reaction, and bridging the distance between the active sites the PCP visits^{21,22}. The A_{sub} uses different surfaces for each of these functions (Extended Data Fig. 9a). In addition, the F domain adds to the long list of partners with which the equally small PCP domain must interact (A, F, C, thioesterase, all tailoring domains), and it performs these functions with overlapping surfaces²⁵ (Extended Data Fig. 9b).

Adapting a formyltransferase has further increased the functionality of NRPSs. The formyl functionality seems to be useful in nonribosomal peptides, as F domains have been incorporated into NRPSs multiple independent times: the F domains in kolossin A synthetase²⁶, anabaenopeptidase synthetases²⁴, the oxazolomycin synthetase²⁷, and other orphan NRPSs and NRPS–polyketide-synthetases arose from a separate fusion event with an FMT and display a different FMT–FMT_{CTD}–C_{partial}–A–PCP domain sequence in their initiation modules⁵. Sampling additional chemical space can lead to novel or improved activity in nonribosomal peptides, which has inspired many bioengineering experiments on NRPSs²⁸ aimed at meeting the dynamic challenges of human health. NRPSs are well placed to be engineered for production of new compounds because their synthetic scheme is

conceptually straightforward, and NRPSs already naturally produce many therapeutics, as well as promising NRPs like teixobactin²⁹ and piperidamycin³⁰, two recently discovered first-in-class compounds with strong antibacterial activity. The structures presented here reveal the interface between the F and A domains and show the interactions that the PCP domain makes in the LgrA initiation module. This knowledge could substantially facilitate our ability to introduce an F domain into a foreign NRPS, and make formylation accessible in NRPS bioengineering.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 August; accepted 2 December 2015.

- Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805–1810 (2004).
- Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nature Chem. Biol.* **11**, 660–670 (2015).
- Hur, G. H., Vickery, C. R. & Burkart, M. D. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat. Prod. Rep.* **29**, 1074–1098 (2012).
- Walsh, C. T. *et al.* Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. *Curr. Opin. Chem. Biol.* **5**, 525–534 (2001).
- Kessler, N., Schuhmann, H., Morneweg, S., Linne, U. & Marahiel, M. A. The linear pentadecapeptide gramicidin is assembled by four multimodular nonribosomal peptide synthetases that comprise 16 modules with 56 catalytic domains. *J. Biol. Chem.* **279**, 7413–7419 (2004).
- Schoenafinger, G., Schracke, N., Linne, U. & Marahiel, M. A. Formylation domain: an essential modifying enzyme for the nonribosomal biosynthesis of linear gramicidin. *J. Am. Chem. Soc.* **128**, 7406–7407 (2006).
- Lawen, A. & Zocher, R. Cyclosporin synthetase. The most complex peptide synthesizing multienzyme polypeptide so far described. *J. Biol. Chem.* **265**, 11355–11360 (1990).
- Robbel, L. & Marahiel, M. A. Daptomycin, a bacterial lipopeptide synthesized by a nonribosomal machinery. *J. Biol. Chem.* **285**, 27501–27508 (2010).
- Konz, D., Klens, A., Schorgendorfer, K. & Marahiel, M. A. The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* **4**, 927–937 (1997).
- Cheng, Y. Q. Deciphering the biosynthetic codes for the potent anti-SARS-CoV cyclodepsipeptide valinomycin in *Streptomyces tsusimaensis* ATCC 15141. *ChemBioChem* **7**, 471–477 (2006).
- Schneider, T. L., Shen, B. & Walsh, C. T. Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation. *Biochemistry* **42**, 9722–9730 (2003).
- Koketsu, K., Minami, A., Watanabe, K., Oguri, H. & Oikawa, H. Pictet-Spenglerase involved in tetrahydroisoquinoline antibiotic biosynthesis. *Curr. Opin. Chem. Biol.* **16**, 142–149 (2012).
- Schmitt, E., Panvert, M., Blanquet, S. & Mechulam, Y. Crystal structure of methionyl-tRNA^{Met} transformylase complexed with the initiator formyl-methionyl-tRNA^{Met}. *EMBO J.* **17**, 6819–6826 (1998).
- Almasy, R. J., Janson, C. A., Kan, C. C. & Hostomska, Z. Structures of apo and complexed *Escherichia coli* glycylamide ribonucleotide transformylase. *Proc. Natl Acad. Sci. USA* **89**, 6114–6118 (1992).
- Thoden, J. B., Gonneau, M. F., Gilbert, M. & Holden, H. M. Structure of a sugar N-formyltransferase from *Campylobacter jejuni*. *Biochemistry* **52**, 6114–6126 (2013).
- Wallace, B. A. Common structural features in gramicidin and other ion channels. *Bioessays* **22**, 227–234 (2000).
- Conti, E., Stachelhaus, T., Marahiel, M. A. & Brick, P. Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *EMBO J.* **16**, 4174–4183 (1997).
- Yonus, H. *et al.* Crystal structure of DltA. Implications for the reaction mechanism of non-ribosomal peptide synthetase adenylation domains. *J. Biol. Chem.* **283**, 32484–32491 (2008).
- Reger, A. S., Wu, R., Dunaway-Mariano, D. & Gulick, A. M. Structural characterization of a 140 degrees domain movement in the two-step reaction catalyzed by 4-chlorobenzoate:CoA ligase. *Biochemistry* **47**, 8016–8025 (2008).
- Gulick, A. M. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem. Biol.* **4**, 811–827 (2009).
- Mitchell, C. A., Shi, C., Aldrich, C. C. & Gulick, A. M. Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains. *Biochemistry* **51**, 3252–3263 (2012).
- Tanovic, A., Samel, S. A., Essen, L. O. & Marahiel, M. A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **321**, 659–663 (2008).
- Liu, Y. & Bruner, S. D. Rational manipulation of carrier-domain geometry in nonribosomal peptide synthetases. *ChemBioChem* **8**, 617–621 (2007).
- Rouhiainen, L. *et al.* Genes encoding synthetases of cyclic depsipeptides, anabaenopeptilides, in *Anabaena* strain 90. *Mol. Microbiol.* **37**, 156–167 (2000).
- Lohman, J. R. *et al.* The crystal structure of Bml1 as a model for nonribosomal peptide synthetase peptidyl carrier proteins. *Proteins* **82**, 1210–1218 (2014).
- Bode, H. B. *et al.* Structure elucidation and activity of kolossin A, the d-/l-pentadecapeptide product of a giant nonribosomal peptide synthetase. *Angew. Chem. Int. Ed. Engl.* **54**, 10352–10355 (2015).
- Zhao, C. *et al.* Oxazolomycin biosynthesis in *Streptomyces albus* JA3453 featuring an “acyltransferase-less” type I polyketide synthase that incorporates two distinct extender units. *J. Biol. Chem.* **285**, 20097–20108 (2010).
- Clardy, J., Fischbach, M. A. & Walsh, C. T. New antibiotics from bacterial natural products. *Nature Biotechnol.* **24**, 1541–1550 (2006).
- Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).
- Hosaka, T. *et al.* Antibacterial discovery in actinomycetes strains with mutations in RNA polymerase or ribosomal protein S12. *Nature Biotechnol.* **27**, 462–464 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank T. Mintya, D. Avizonis, M.-C. Tang and A. Furtos for experimental support, R. Zamboni, J. Collucci and K. Guerard for small molecule synthesis assistance; J. Jiang, D. Alonzo and D. Rodinov for experimental advice and assistance; S. Labuik and P. Grochulski (Canadian Light Source) for diffraction data collection; R. Gillian (CHESS SAXS beamline G1), M. Pillon and A. Guarne for SAXS help; K. Bloudoff, M. Tarry, A. Haque and A. Beghuis for helpful discussions; and J. Pelletier, N. Rogerson and A. Nahvi for critical reading of the manuscript. This work was supported by CIHR grant 106615, a HFSP CDA and a Canada Research Chair in Macromolecular Machines to T.M.S. J.M.R. is supported by an NSERC Alexander Graham Bell studentship, and M.N.A. by a studentship from the CIHR Training Grant in Chemical Biology.

Author Contributions J.M.R. and M.N.A. cloned, expressed, purified, assayed, crystallized and determined the structure of LgrA F–A and F–A–PCP proteins. J.M.R. performed the CoA syntheses and prepared figures. P.M.H. performed the bioinformatics analyses. T.M.S. designed the study, analysed data and wrote the manuscript with input from the other authors.

Author Information Crystallography data have been deposited in the Protein Data Bank under accession codes 5ES5, 5ES6, 5ES7, 5ES8 and 5ES9. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.M.S. (martin.schmeing@mcgill.ca).

METHODS

Data reporting. No statistical methods were used to determine sample size.

Cloning of linear gramicidin synthetase initiation module constructs.

Genomic DNA was isolated from *Brevibacillus parabrevis* ATCC 8185 (Cedarlane Laboratories) using a GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich). Gene constructs comprising F and A domains (F–A) and all three domains (F–A–PCP) were amplified by PCR from the *lgrA* gene using the following primers, designed using sequence alignment with A and PCP domains of a known structure and the study of Marahiel and co-workers^{5,6}. FA_fwd: 5′-AATCA TCCATGGGAAGAATACTATTCCCTAACACATTTATGAGCAAAG-3′; FA_rev: 5′-AATCATCTCGAGTTACGCATCGGCCTGCACGTCT-3′; FAT_fwd: 5′-TGACTACCATGGGAGAATACTATTCCCTAACACATTTATGAGC-3′; FAT_rev: 5′-CGTTGAGCGGCCGCTTGCTCCGTAAGCAGACGTTT-3′. PCR product for F–A–PCP was digested using NcoI and NotI (New England Biolabs) and ligated into a pET21-derived vector containing an N-terminal octa-histidine tag with a tobacco etch virus (TEV) protease cleavage site. The PCR product for F–A–PCP was cloned between NcoI and NotI restriction sites into a pET21-derived vector containing an N-terminal TEV cleavable octa-histidine tag and a C-terminal TEV cleavable calmodulin binding peptide (CBP) tag.

Expression and purification of proteins. The F–A protein was expressed in *Escherichia coli* BL21 (DE3) cells. A 10 ml aliquot of overnight culture was used to inoculate 1 l of lysogeny broth (LB) medium supplemented with 350 µg ml^{−1} kanamycin. The culture was grown at 37 °C to an optical density (OD₆₀₀) of 0.6, before inducing protein expression using 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and reducing the temperature to 16 °C for 18 h. Cells were collected by centrifugation at 4 °C and resuspended in nickel binding buffer (2 mM imidazole, 150 mM NaCl, 0.25 mM tris-(2-carboxyethyl)phosphine (TCEP), 50 mM Tris-HCl (pH 7.0)). The cells were lysed by sonication on ice and centrifuged for 30 min at 20,000g at 4 °C. Clarified lysate was loaded onto a HiTrap IMAC FF column (GE Healthcare). F–A protein was eluted using a gradient of 2–250 mM imidazole. Fractions containing F–A were pooled, diluted tenfold with ion exchange binding buffer (0.25 mM TCEP, 20 mM Tris, pH 8.0), loaded onto a HiTrap Q HP column and eluted using a gradient to 100% elution buffer (1 M NaCl, 0.25 mM TCEP, 20 mM Tris-HCl (pH 8.0)). The eluted protein was concentrated using a 10K MWCO Amicon Ultra-15 filtration unit (EMD Millipore) and subjected to gel filtration chromatography using a HiLoad 16/600 Superdex 200 column (GE Healthcare) equilibrated with S200 buffer (150 mM NaCl, 0.25 mM TCEP, 20 mM Tris (pH 7.0)). Protein purity was confirmed using SDS–PAGE and native PAGE. Pure F–A was concentrated in storage buffer (25% glycerol, 150 mM NaCl, 0.25 mM TCEP, 20 mM Tris (pH 7.0)), flash-frozen with liquid nitrogen and stored at −80 °C for later use.

F–A–PCP was expressed in *E. coli* BL21 EntD-(DE3) cells using the same protocol as above. Cells were pelleted, resuspended in CBP binding buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2 mM imidazole (pH 8.0), 2 mM CaCl₂, 2 mM β-mercaptoethanol (βME) and 0.1 mM phenylmethanesulfonyl fluoride (PMSF)), sonicated and clarified by centrifugation for 30 min at 20,000g at 4 °C. Clarified lysate was loaded onto a 30 ml calmodulin sepharose 4B column (GE Healthcare). F–A–PCP was eluted with elution buffer (25 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2 mM EGTA, 2 mM βME and 0.1 mM PMSF). Protein was dialysed against binding buffer for a minimum of 4 h before being loaded onto a 5 ml HiTrap IMAC FF column (GE Healthcare) charged with Ni²⁺ and equilibrated in nickel binding buffer. F–A–PCP was eluted using a 60 ml gradient of 0–250 mM imidazole. Fractions containing F–A–PCP were pooled and affinity tags were removed by cleavage with TEV protease at room temperature overnight using a 1:4 mg ratio of TEV to F–A–PCP. Cleaved F–A–PCP was passed back over the nickel and calmodulin affinity columns, with the flow-through collected, concentrated and applied to a HiLoad 16/600 Superdex 200 (GE Healthcare) in S200 buffer. Pure F–A–PCP was concentrated to 5.0 mg ml^{−1} in storage buffer, flash-frozen in liquid nitrogen and stored at −80 °C.

Substrate syntheses. Amino-coenzyme A (NH-CoA)²³ was prepared enzymatically starting from amino-pantetheine (WuXi AppTec) using a previously published protocol³¹ with the following modifications: one-pot synthesis was carried out at pH 9.0; the amounts of DPCCK and ATP were doubled to 9.8 mg and 30 mM, respectively; and the enzymes were removed using a 10K MWCO Amicon Ultra-15 filtration unit (EMD Millipore). An ATP regeneration system using 0.1 mg ml^{−1} pyrophosphatase (Roche), 30 mM phosphoenolpyruvate, and 0.1 mg ml^{−1} pyruvate kinase (Roche) was also included. The filtrate containing NH-CoA was purified on a preparative reverse-phase C18 HPLC (35 ml min^{−1}; 0–4 min, 0% B; 4–9 min, 0–98% B, where A is 0.1% trifluoroacetic acid (TFA; Sigma-Aldrich) in H₂O and B is 0.1% in acetonitrile (ACN; Sigma-Aldrich)). NH-CoA was eluted at 7 min and lyophilized to dryness.

Valine-amino-CoA (Val-NH-CoA)²³ was synthesized by coupling 1 molar equivalent of NH-CoA with 8 molar equivalents of *tert*-butoxycarbonyl-L-valine-*N*-hydroxysuccinimide ester (Boc-Val-OSu; Sigma-Aldrich) in *N,N*-dimethylformamide (DMF; Sigma-Aldrich) with 4 molar equivalents of *N,N*-diisopropylethylamine (DIPEA; Sigma-Aldrich) overnight with stirring. Boc-Val-NH-CoA was purified using the above chromatographic profile and lyophilized to dryness, then deprotected using 1.5 ml 95% TFA, 2.5% H₂O and 2.5% triisopropylsilane (TIPS; Sigma-Aldrich). The deprotection mix was agitated for 2 h at 25 °C in a thermomixer at 700 r.p.m. before being transferred to 20 ml ice-cold diethyl ether and incubated at −20 °C for 2 h. The solution was centrifuged and the pellet was redissolved in 5% aqueous ACN solution and purified with the same protocol as NH-CoA. Compound identity was verified by mass spectrometry and nuclear magnetic resonance (NMR) (Supplementary Data 1).

Loading phosphopantetheinylates on the PCP domain. Unmodified F–A–PCP was converted to Val-NH-F–A–PCP by incubating 25 µM apo-F–A–PCP with 5 µM of the promiscuous phosphopantetheinyl transferase Sfp, 0.25 mM Val-NH-CoA, 10 mM MgCl₂ and 25 mM Tris (pH 7.0) for a minimum of 4 h at 25 °C. To remove Sfp for subsequent crystallization trials, the reaction mix was loaded onto a Superdex S75 10/300 GL (GE Healthcare Life Sciences) equilibrated in 25 mM Tris (pH 7.5), 150 mM NaCl and 2 mM βME.

SAXS. Inline size exclusion chromatography with small-angle X-ray diffraction (SEC–SAXS) data was collected on the G1 beamline at the Macromolecular Diffraction Facility at the Cornell High Energy Synchrotron Source^{32,33} at 9.963 keV (1.244 Å) at 7.89 × 10¹¹ photons s^{−1}. The X-ray beam was collimated to 250 × 250 µm and the sample cell path length was 2 mm. The G1 beamline was outfitted with a GE AKTA purifier with a GE Superdex 200 5/150 GL column and 50 µl sample loop. The column was equilibrated in 25 mM Tris (pH 7.5), 150 mM NaCl and 2 mM βME and the samples were centrifuged for 10 min before sample injection. Images were recorded on a Pilatus 100K-s detector and normalized using beam stop photodiode counts. F–A–PCP eluted in a single monomeric peak and eleven peak exposures were averaged using BioXTAS RAW software³⁴. A buffer scattering curve was created by averaging the first eleven exposures after injection, and this scattering curve was subtracted from the F–A–PCP scattering curve to yield the corrected scattering curve for F–A–PCP. *Ab initio* models were generated by first creating pairwise distribution functions (*P(r)*) with GNOM³⁵, leading to twenty independent bead models produced by DAMMIF³⁶. Models were aligned, averaged, and filtered using DAMAVER³⁷ assuming *P1* symmetry. All DAMMIF models were included in the final DAMAVER model. They had a mean normalized spatial discrepancy (NSD) value of 0.82 ± 0.052. CRY SOL³⁸ was used to check how well the final model fit with our crystal structures. Flexibility was analysed using EOM^{39,40}, whereby crystal structures of F, A_{core}, A_{sub} and the PCP were used to generate a pool of 10,000 models.

Crystallography. To obtain the crystal structures described in this study, genes from four species, of up to four domain constructs each (F, F–A, F–A(ΔA_{sub}) and F–A–PCP), were cloned and assayed for heterologous expression. Purification was performed for all highly expressing proteins and crystallization trials were performed, including trials using protein with affinity tags removed or retained, and in the presence or absence of a variety of ligands (ATP, AMPcPP, AMP, valine, THF, N⁵-fTHF, phosphopantetheine, valine amino phosphopantetheine, valine vinyl sulfonamide adenylate, dead-end THF analogue). Up to 4,032 crystallization conditions were assayed per protein sample, and gave a total of ~50 ‘hits’, 6 of which were successfully optimized to allow structure determination. Together, 4 of these crystal structures (F–A in crystals of space group P4₁2₁2, F–A–PCP in R3:H, F–A–PCP–PPE–NH–Val in P2₁ and F–A–PCP–PPE in P3₂2), plus an additional structure including ligands soaked into F–A P4₁2₁2 crystals, captured the states that represent every major step of the assembly-line synthesis in the *LgrA* initiation module and are presented here.

The final crystallization conditions were optimized in 24-well sitting drop plates, with 2 µl protein sample plus 2 µl reservoir solution in the drop and a 500 µl reservoir volume, and are as follows. ‘F–A’ and ‘F–A soak’: protein *LgrA* F–A (10 mg ml^{−1}) was crystallized using a precipitant solution of 2 M Na-formate, 0.1 M sodium acetate (pH 5.3) into space group P4₁2₁2. ‘F–A–PCP’ (open and closed states): protein *LgrA* F–A–PCP (5 mg ml^{−1}) was crystallized using a precipitant solution of 0.92 M AmSO₄, 0.1 M bis-Tris (pH 5.5), 1% polyethylene glycol (PEG) 3350 into space group R3:H. ‘F–A–PCP–NH–Val’ (thiolation state): protein F–A–PCP–PPE–NH–Val (4.7 mg ml^{−1}) was crystallized using a precipitant solution of 12% PEG 20,000, 0.1 M MES buffer (pH 6.7) into space group P2₁. ‘F–A–PCP–PPE’ (formylation state): protein F–A–PCP–PPE (5.5 mg ml^{−1}) was crystallized using a precipitant solution of 1 M AmSO₄, 0.1 M bis-Tris (pH 5.5), 3% PEG 3350 into space group P3₂2.

Solutions of mother liquor with increasing amounts of glycerol (5%, 10%, 25%) were used to replace the drop solution for cryoprotection. For soaking with the

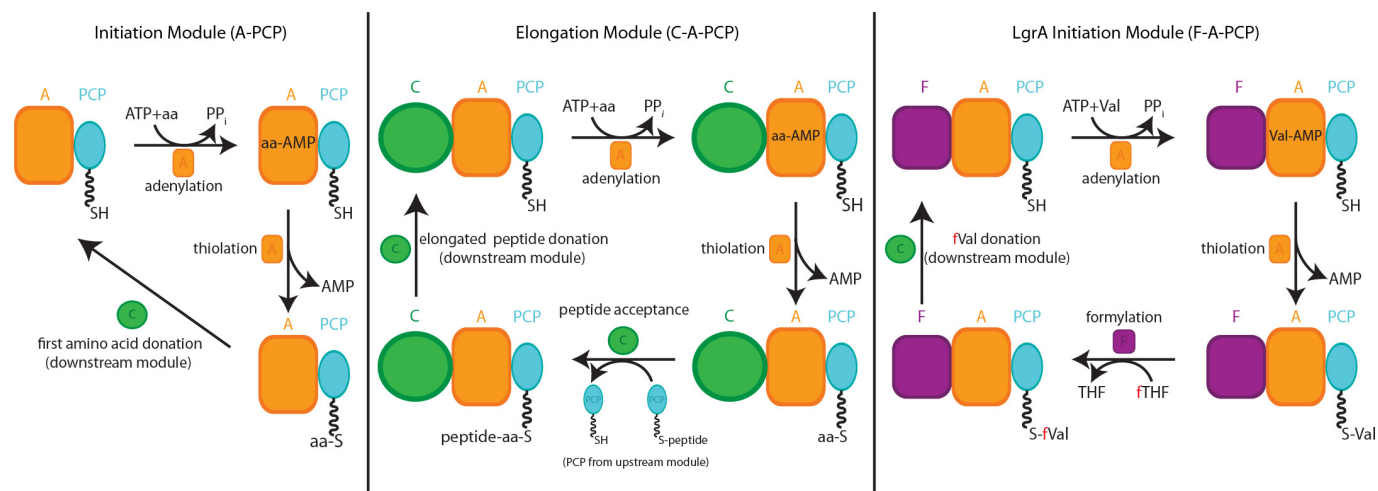
N⁵-fTHF, valine and AMPcPP, 10 mM of each was included in the final cryoprotection solution and incubated for 30 min. (LgrA uses commercially available N⁵-fTHF in addition to its natural substrate, N¹⁰-fTHF⁶.) Crystals were flash-cooled in liquid nitrogen and diffraction data sets collected at 200 K using beamline 8 of the CMCF at the Canadian Light Source ($\lambda = 0.979$ Å) in Saskatoon, Canada.

All data sets were integrated and scaled using the programs HKL-2000 (ref. 41) and iMosflm⁴². Structure determination of F–A in the $P_4_2_1_2$ space group was performed by molecular replacement using a search model of the A domain from gramicidin Soviet synthetase¹⁷ (note that linear gramicidin and gramicidin Soviet are made by different NRPSs) with the A_{sub} subdomain removed and side chains trimmed to the β -carbon, in the program Phaser⁴³. Density for the F domain was visible in the resulting maps. Iterative building in the program COOT⁴⁴ and refinement in the program Phenix⁴⁵ produced the final F–A structure. This structure was then used as a search model to determine the structure of F–A–PCP in space groups $P_3_2_2$, $R_3:H$, and P_2_1 by molecular replacement using the program Phaser, followed by iterative building in the program COOT and refinement in the programs Phenix and CNS⁴⁶. The highest resolution shell CC* values are: $P_4_2_1_2$, 0.845; $P_4_2_1_2$ (soak), 0.897; $P_3_2_2$, 0.883; $R_3:H$, 0.822; and P_2_1 , 0.822. The quoted resolution of each structure represents the half-data set correlation coefficient (CC 1/2) of the diffraction data^{42,47}.

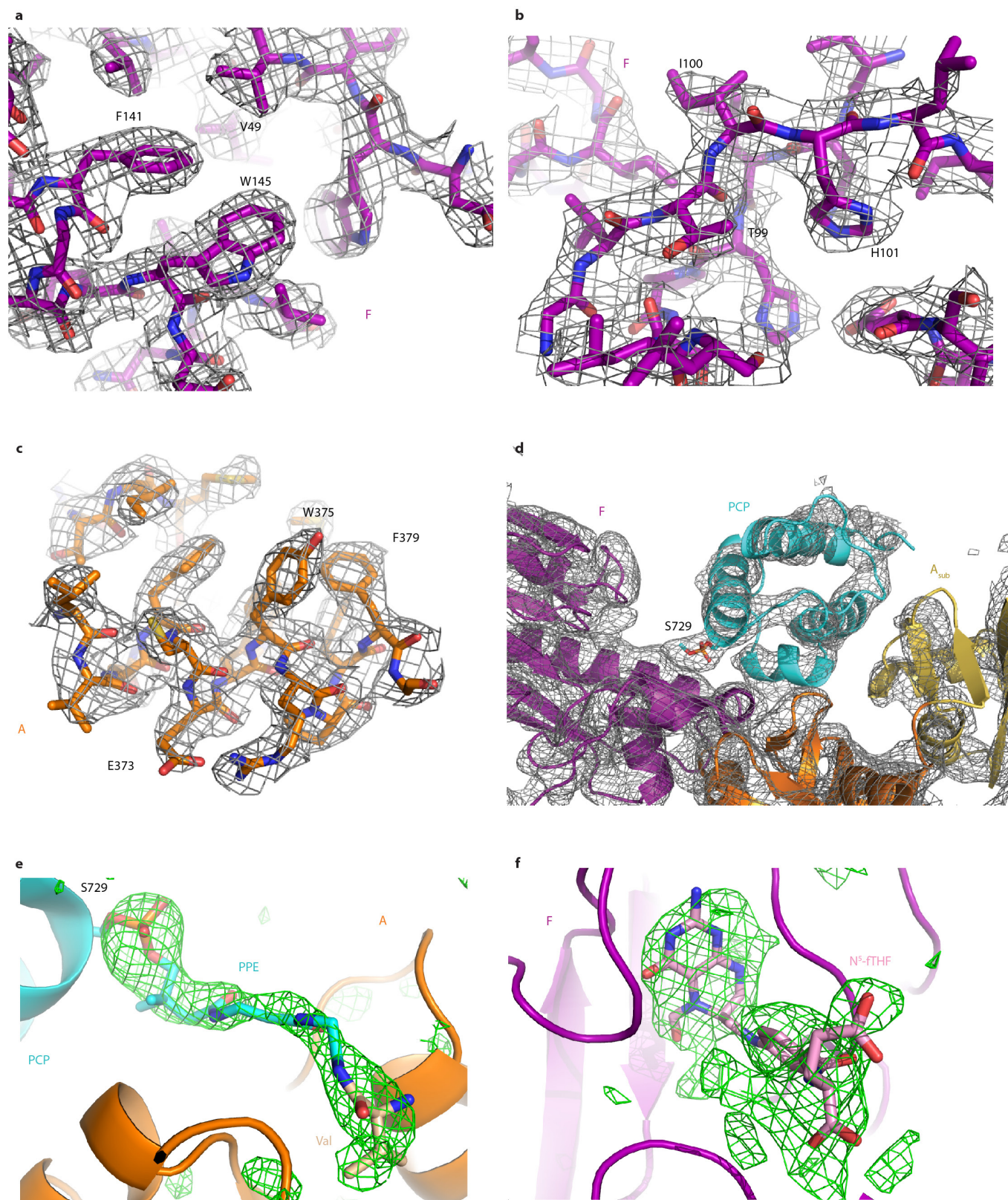
Bioinformatics. Multiple sequence alignments (MSAs) were constructed using Clustal Omega⁴⁸ (<http://www.ebi.ac.uk/Tools/msa/clustalo>) and PROMALS3D⁴⁹ (<http://proddata.swmed.edu/promals3d>), following database searches using BLAST⁵⁰ (<http://ncbi.nlm.nih.gov/blast>). MSAs were drawn/edited using Jalview⁵¹ (<http://www.jalview.org>). PHYLP (<http://evolution.genetics.washington.edu>) was used to make neighbour-joining trees bootstrapped with 100 replicates, and FigTree (<http://tree.bio.ed.ac.uk>) was used to draw them. WebLogo⁵² (<http://weblogo.berkeley.edu>) was used to draw sequence logos of residue groupings of interest. AmiGO⁵³ (<http://amigo.geneontology.org>) was used to check for experimentally characterized proteins.

Analysis of synthesized Val–NH–CoA. Val–NH–CoA was verified by both mass spectrometry (calculated m/z [MH^+]: 850.2304; measured m/z [MH^+]: 850.2299) and ¹H NMR [¹H NMR (600 MHz, H₂O) δ 8.69 (d, $J = 16.2$ Hz, 1H), 8.46 (s, 1H), 6.25 (d, $J = 5.9$ Hz, 1H), 4.97–4.89 (m, 1H) 4.75–4.73 (s, 1H), 4.65–4.58 (m, 1H), 4.30–4.21 (m, 2H), 4.04 (s, 1H), 3.91–3.82 (d, $J = 5.7$ Hz, 1H), 3.75 (dd, $J = 8.8$ Hz, 1H), 3.68–3.61 (d, $J = 5.9$ Hz, 1H), 3.55–3.42 (m, 3H), 3.42–3.23 (m, 4H), 2.48 (m, 2H), 2.21–2.14 (m, 1H), 1.03–0.99 (m, 7H), 0.87 (s, 3H), 0.44 (s, 3H)].

31. Nazi, I., Koteva, K. P. & Wright, G. D. One-pot chemoenzymatic preparation of coenzyme A analogues. *Anal. Biochem.* **324**, 100–105 (2004).
32. Acerbo, A. S., Cook, M. J. & Gillilan, R. E. Upgrade of MacCHESS facility for X-ray scattering of biological macromolecules in solution. *J. Synchrotron Radiat.* **22**, 180–186 (2015).
33. Skou, S., Gillilan, R. E. & Ando, N. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nature Protocols* **9**, 1727–1739 (2014).
34. Nielsen, S. S. *et al.* BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *J. Appl. Crystallogr.* **42**, 959–964 (2009).
35. Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
36. Franke, D. & Svergun, D. I. DAMMIF, a program for rapidab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).
37. Volkov, V. V. & Svergun, D. I. Uniqueness of ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864 (2003).
38. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOLE – A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
39. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCr J* **2**, 207–217 (2015).
40. Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).
41. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
42. Leslie, A. G. W. & Powell, H. R. Processing diffraction data with MOSFLM. *Nato Sci Ser II Math* **245**, 41–51 (2007).
43. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
44. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
45. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
46. Brunger, A. T. Version 1.2 of the crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
47. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013).
48. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
49. Pei, J. & Grishin, N. V. PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. *Methods Mol. Biol.* **1079**, 263–271 (2014).
50. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
51. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. & Barton, G. J. Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
52. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
53. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. *Nature Genet.* **25**, 25–29 (2000).
54. Allen, C. L. & Gulick, A. M. Structural and bioinformatic characterization of an *Acinetobacter baumannii* type II carrier protein. *Acta Crystallogr. D* **70**, 1718–1725 (2014).
55. Lai, J. R., Fischbach, M. A., Liu, D. R. & Walsh, C. T. A protein interaction surface in nonribosomal peptide synthesis mapped by combinatorial mutagenesis and selection. *Proc. Natl Acad. Sci. USA* **103**, 5314–5319 (2006).
56. Jaremko, M. J., Lee, D. J., Opella, S. J. & Burkart, M. D. Structure and substrate sequestration in the pyoluteorin type II peptidyl carrier protein PtlL. *J. Am. Chem. Soc.* **137**, 11546–11549 (2015).
57. Goodrich, A. C., Harden, B. J. & Frueh, D. P. Solution structure of a nonribosomal peptide synthetase carrier protein loaded with its substrate reveals transient, well-defined contacts. *J. Am. Chem. Soc.* **137**, 12100–12109 (2015).

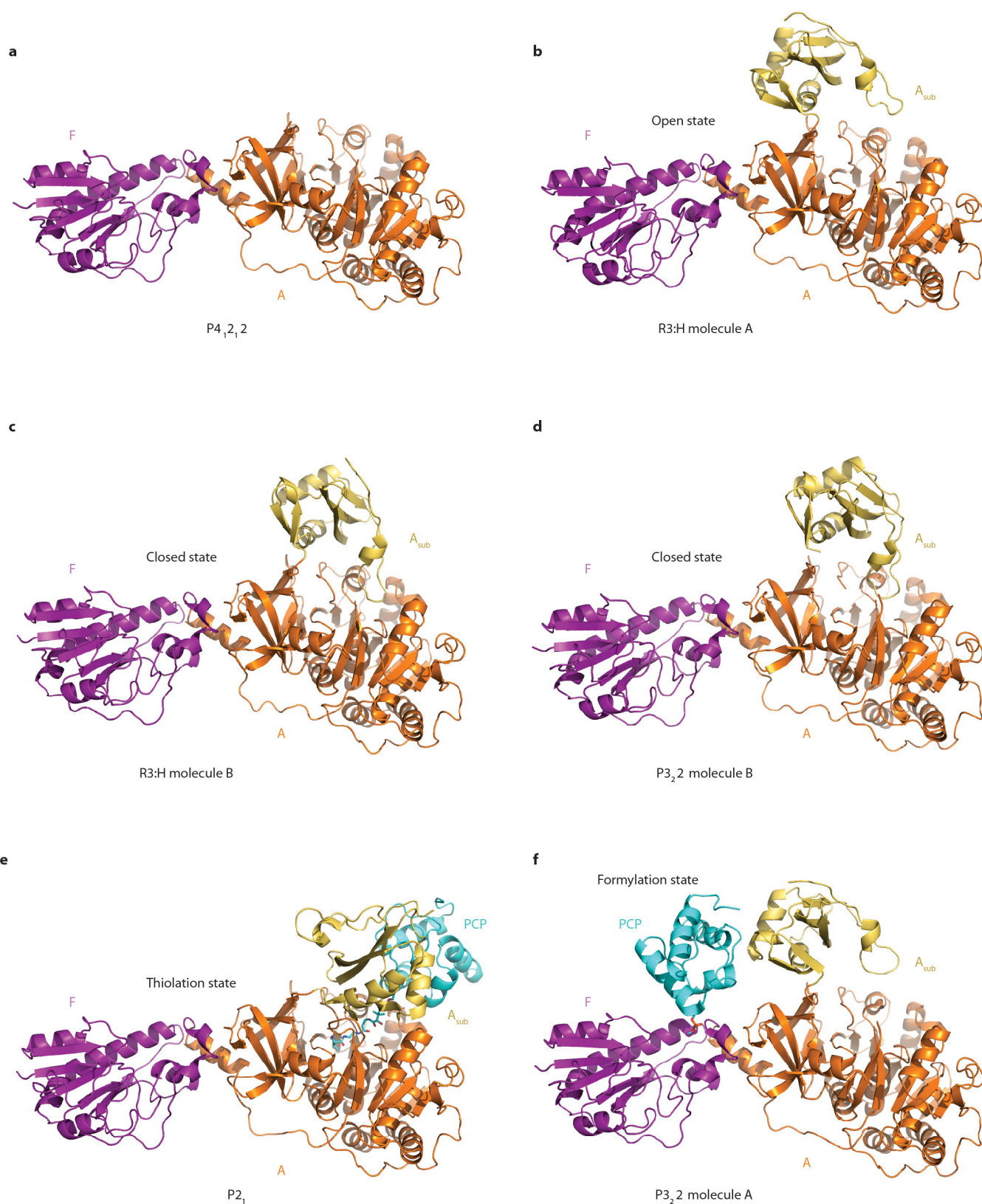


Extended Data Figure 1 | Synthetic cycles in canonical initiation, canonical elongation and LgrA initiation modules. Schematic diagrams comparing the synthetic cycle in canonical initiation and elongation modules with that in the LgrA initiation module.



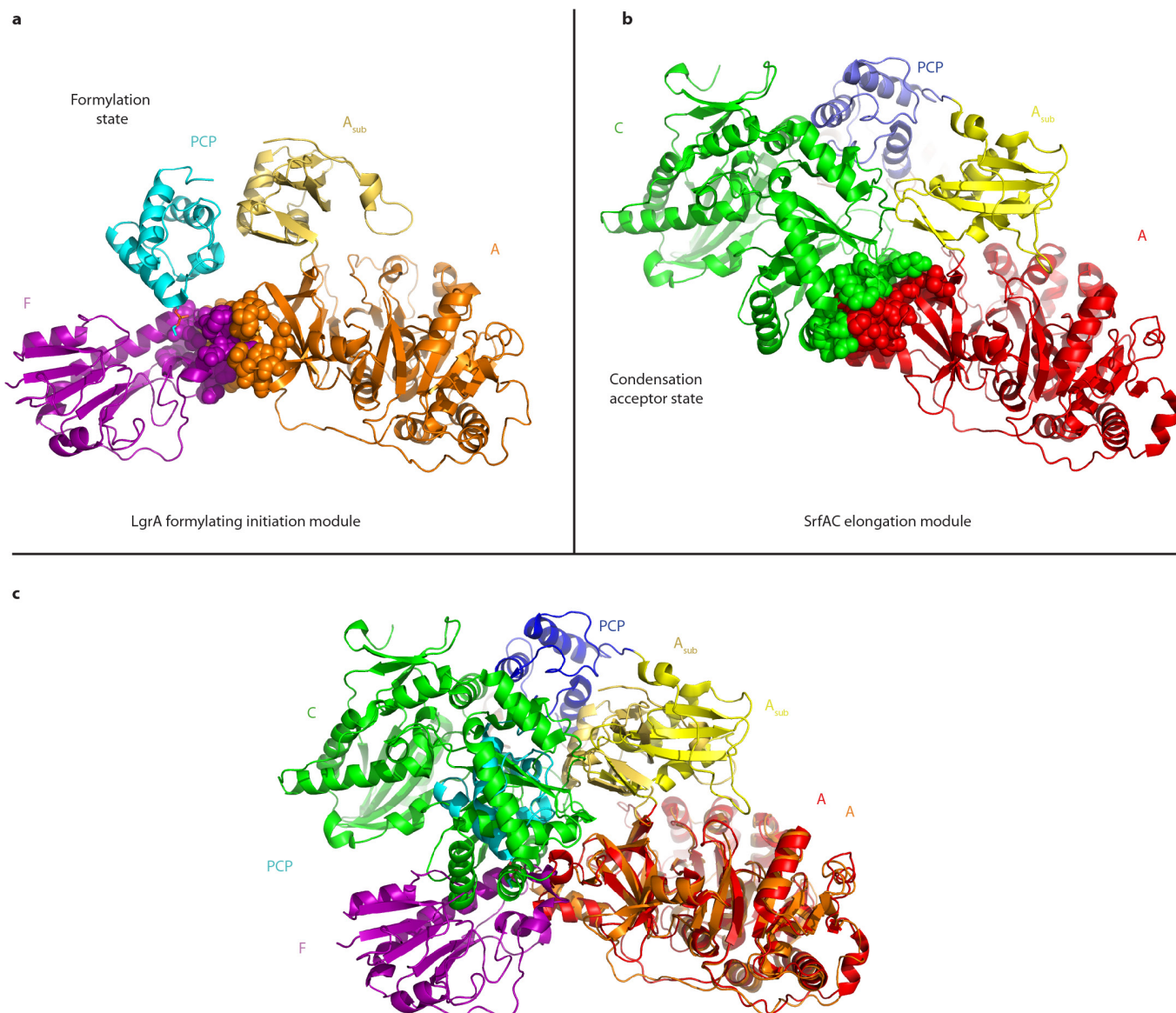
Extended Data Figure 2 | Representative electron density. a–d, $2F_o - F_c$ density maps for protein in $P4_12_12$ (a), $R3:H$ (b), $P2_1$ (c) and $P3_22$ (d) crystal forms contoured at 1σ . e, f, Unbiased $F_o - F_c$ density maps for the

PPE-NH-Val arm in the $P2_1$ (thiolation state) (e), and a $P4_12_12$ crystal soaked with N5-fTHF, AMPcPP and valine contoured at 2.5σ (f).



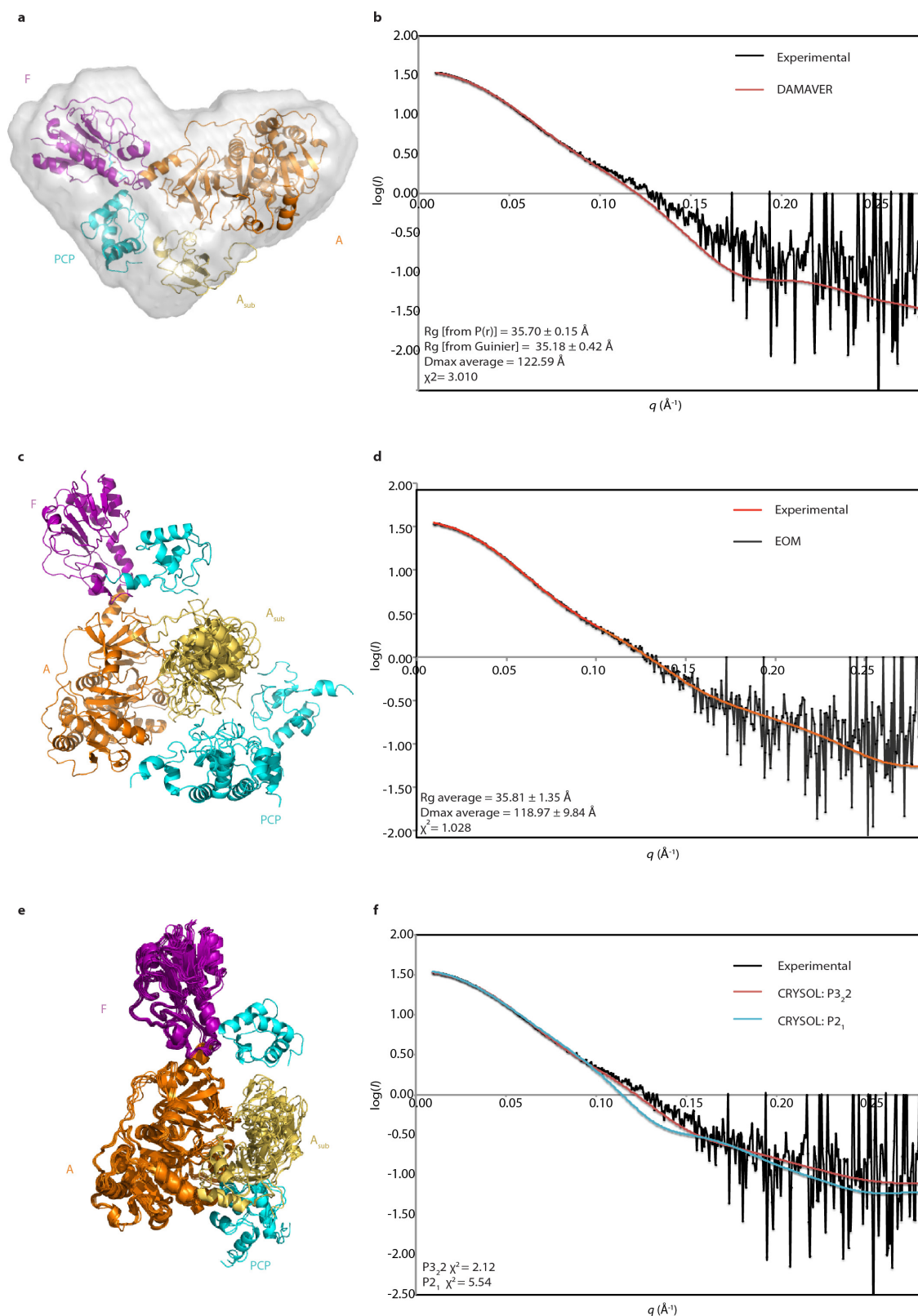
Extended Data Figure 3 | Crystal structures of the initiation module of linear gramicidin synthetase. a–f, Models of F–A (A_{sub} disordered) (a), F–A–PCP (PCP disordered) (b–d) and F–A–PCP from the four independent crystals structures determined (e, f). The crystal with space

group $P3_2$ diffracted anisotropically to ~ 3.8 Å resolution, but the other higher resolution structures enabled the building of high quality models shown in d and f.



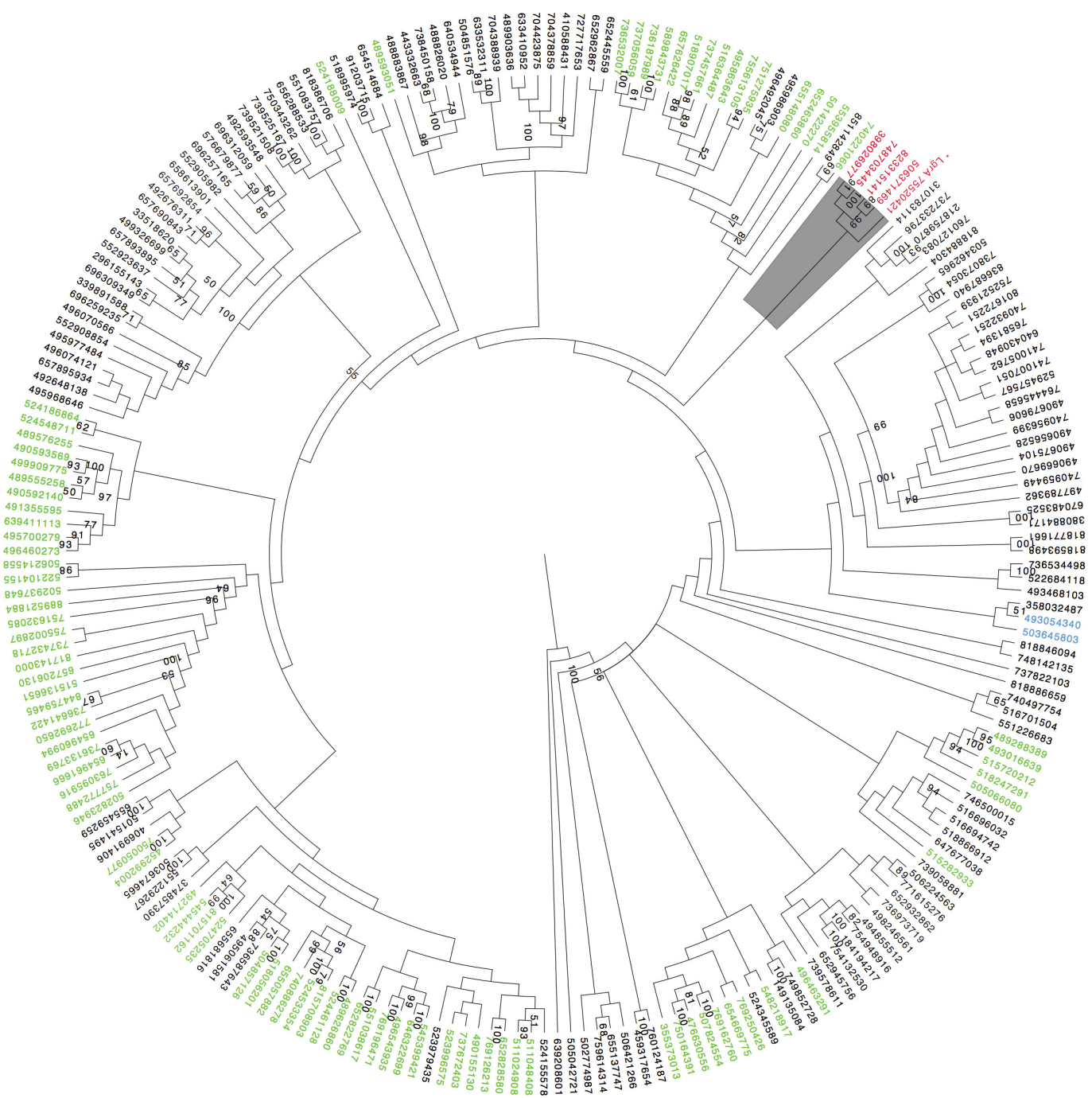
Extended Data Figure 4 | Comparison between the LgrA initiation module and the SrfAC termination module. **a, b,** The LgrA initiation module in the formylation state (**a**) and the termination module of surfactin synthase subunit 3 (SrfAC)²² (**b**) in the state where aminoacyl-PCP would be positioned to act as an acceptor substrate in the condensation reaction (PPE arm not present). The F and C domains are each positioned directly N-terminal of their A domains and bury similar amounts of A domain surface area (829 Å² and 903 Å²; contributing residues shown in spheres), each forming 'stable platforms'²².

Both modules use very large movements of their PCP and A_{sub} domains to bring the aminoacyl-PCP of the module to distant active sites to act as the acceptor substrate in an amide bond forming reaction. **c,** However, the F–A and C–A interfaces are distinct, and, if the A domains are superimposed, the F and C domains are only partially overlapping. This places their active sites in dissimilar locations, necessitating that A_{sub} and PCP assume different positions to deliver their substrate. The PCP domain in the formylation state completely overlaps with the position of the C domain.



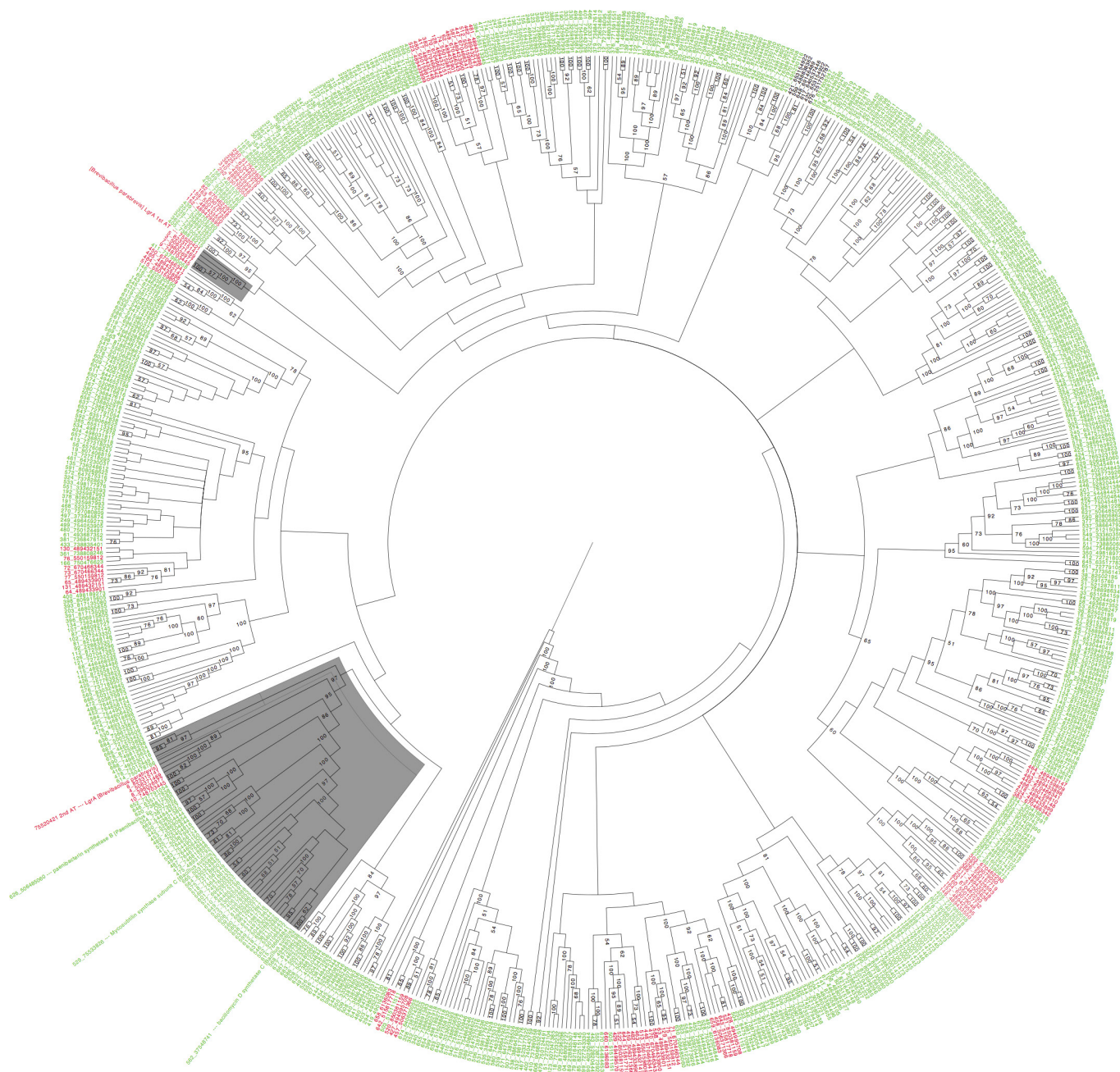
Extended Data Figure 5 | Small-angle X-ray scattering analysis of F-A-PCP. **a**, The crystal structure in the formylation state is shown superposed on the averaged filtered *ab initio* small-angle X-ray scattering model generated with DAMAVER³⁷, with a NSD value of 0.819 ± 0.052 . **b**, The calculated scattering curve for the DAMAVER is overlaid with the experimental scattering with $\chi^2 = 3.010$, where I represents scattering intensity and q is equivalent to $4\pi\sin(\theta)/\lambda$. **c**, To understand the flexibility of F-A-PCP better, EOM³⁹ was performed and generated five different ensembles. The ensemble resembling the formylation state structure represented over 60% of the optimized models generated,

while the remaining <40% resembled the thiolation state structure. **d**, The calculated scattering of the EOM model has a $\chi^2 = 1.028$, which demonstrates that F-A-PCP has flexibility. The data are consistent with extreme flexibility for A_{sub} and PCP domains, and limited flexibility in F- A_{core} . **e**, All independent molecules from the crystal structures were overlaid to further illustrate the flexibility of the system. **f**, CRYSOLE³⁸ was used to generate predicted scattering curves for the formylation state and thiolation state crystal structures with $\chi^2 = 2.12$ and $\chi^2 = 5.54$, respectively.



Extended Data Figure 6 | Neighbour-joining tree of LgrA F domain and homologues. This neighbour-joining tree of the LgrA F domain and homologues was made using PHYLIP (<http://evolution.genetics.washington.edu>) based on an initial Clustal Omega⁴⁸ alignment of the closest 220 homologues of the LgrA F domain (Blast⁵⁰ BLAST E-value $< 1 \times 10^{-14}$). The most similar formyltransferases to the F domain share ~45% identity, and all of these 220 formyltransferases have only inferred function. The tree was drawn using the program FigTree (<http://tree.bio.ed.ac.uk>). The sequences are named with their GenInfo Identifier (GI)

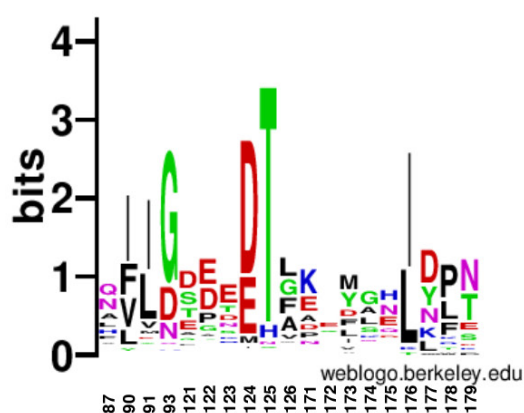
numbers. Colouring: red, *Brevibacilli*; green, other Firmicutes; black, other bacteria; blue, Archaea. The clade of the LgrA F domain is highlighted in grey. Only nodes with bootstraps of $> 50\%$ are shown. Several horizontal transfer events are evident where Firmicute and non-Firmicute proteins cluster together with high bootstrap values (for example, $> 70\%$). The several horizontal transfer events of formyltransferase domains between Firmicutes and other bacterial groups suggest the LgrA F domain likely originated from horizontal transfer.



Extended Data Figure 7 | Neighbour-joining tree of LgrA A-PCP and homologues. This neighbour-joining tree of LgrA A-PCP didomains and homologues was made for the 500 closest homologues (BLAST E-value $<1 \times 10^{-4}$). The sequences are named with their GI accession codes. Colouring: red, *Brevibacilli*; green, other Firmicutes; black, other bacteria. The significant clades of the LgrA A-PCP domains are highlighted in grey.

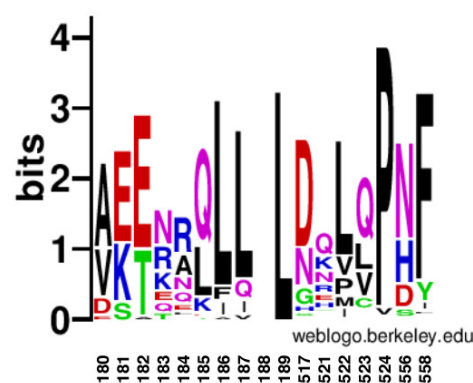
Only nodes with bootstraps of $>50\%$ are shown. Three functionally characterized homologues of LgrA that are shown to be directly related are labelled. The A-PCP portion of the initiation module is quite divergent, but the second module of LgrA clearly shares a common origin with functionally characterized NRPSs in Bacilli and other Firmicutes.

a, FT domain residues interacting with the A domain

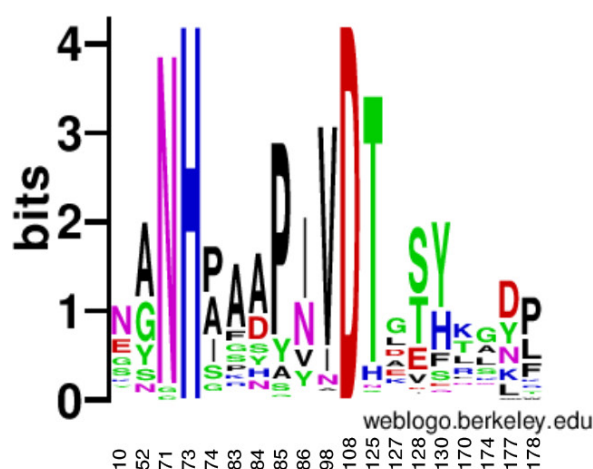


FVWEADEDTLDFYKNLNMT
FIWEDDEDTLDFYKNLNMT
FIWEDDEDTLDFYKNLNMT
FIWEDEEDTLDFYKNLNMT
FIWEDDEDTLDFYKNLNMT

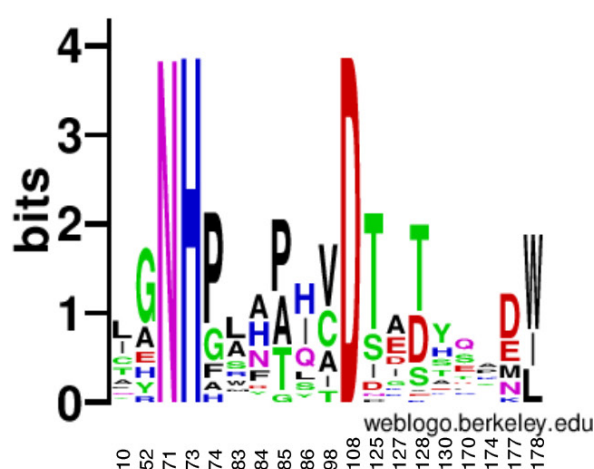
b, A domain residues interacting with the FT domain



TVRELLALKRDALQPHQ
TVKELLALKRDALQPHQ
TVKELLALKRDALQPHQ
TVKELLALKRDALQPHQ
TVKELLALKRDALQPHQ



c, Closest 240 homologues of the F domain



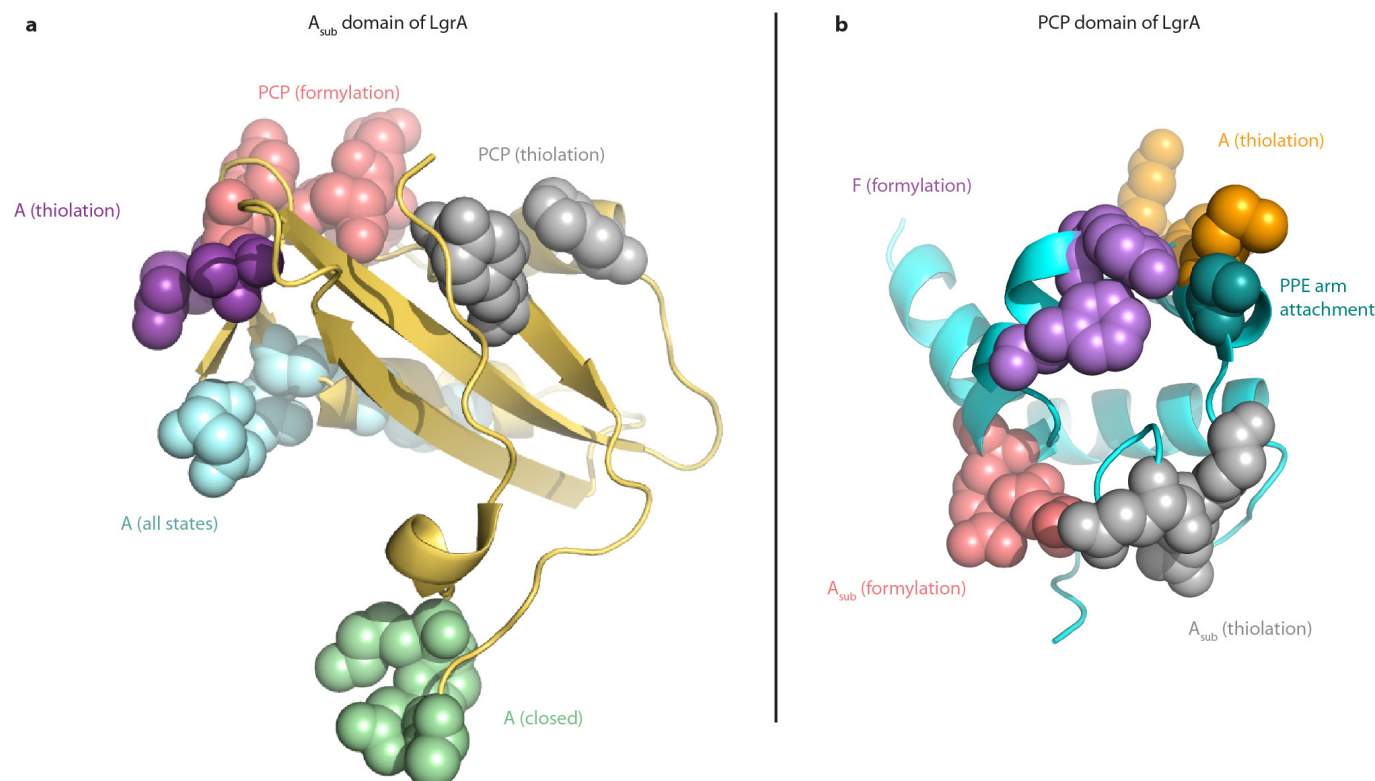
d, Formyltransferases of known function

e, Consensuses of formyltransferase sequences

F	A	N	HT	RDPV	V	D	T	ID	Y	R	K	NM	5 Brev F-domains
10	52	71	73	83	98	108	125	127	130	170	174	177	
I	—	N	H	WF	—	V	D	T	KD	Y	—	—	sugar FTs
L	A	N	HG	AAP	I	D	T	A	Y	S	A	DW	met-tRNA FTs
A	G	N	HP	LHAH	C	D	T	ET	A	Q	P	—	purine FTs

Extended Data Figure 8 | Conservation and variation of residues involved in the interaction interfaces. **a**, **b**, Sequence logos made using the WebLogo server (<http://weblogo.berkeley.edu>)⁵² show conservation and variation as found in multiple sequence alignments of F domain residues that interact with the A domain (**a**) and A domain residues that interact with the F domain (**b**). Below each logo are the corresponding residues in the LgrA proteins from the five *Brevibacillus* species, with the crystallized LgrA on the first line. FT, formyltransferase. **c**, **d**, Sequence

logos indicate the conservation and variation in F domain residues involved in binding and interaction with PCP-PPE-Val across the closest 240 homologues of LgrA (**c**) and all of the functionally or structurally characterized formyltransferase proteins (**d**) (reduced for redundancy so that no two sequences have >50% sequence identity). **e**, Consensus sequences for the five *Brevibacillus* LgrA homologues and for the formyltransferases of known structure for each of three formyltransferase types. Catalytic residues are His73, Asn71 and Asp108.



Extended Data Figure 9 | Interaction surfaces in PCP and A_{sub} domains.

a, b, The A_{sub} (**a**) and PCP^{2,25,54} (**b**) domains must maximize the use of their limited surfaces to interact with their many binding partners. Shown are the surfaces observed in this study, and many excellent previous studies have also documented interaction surfaces biochemically or structurally. This includes, for example, the equivalent of PCP domain residues Met249, Phe264 and Ala268, which are required for interaction with the C domain in the acceptor site⁵⁵ and form hydrophobic interactions with the C domain²² in a very similar manner and using an overlapping surface, as the PCP domain does to interact with the F domain. Furthermore,

partially overlapping surfaces in PCP domains have been proposed to interact with their (acyl-)PPE arm to protect thioester intermediates⁵⁶ or to promote binding to the appropriate partner domain⁵⁷. These interactions might occur during PCP domain transit, but they would have to be broken before productive binding to partner domains. Several of these PPE interactions are incompatible with the productive domain–domain interactions⁵⁷, and in catalytic configurations seen here and previously, the PPE arms extend into the partner domain and make little contact with the PCP domain.

Extended Data Table 1 | Crystallographic statistics

	F-A	F-A-PCP	F-A-PCP-PPE-NH-Val	F-A-PCP-PPE	F-A soak
Data collection					
Space group	P4 ₁ 2 ₁ 2	R3:H	P2 ₁	P3 ₂ 2	P4 ₁ 2 ₁ 2
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	161.3, 161.3, 138.2	278.7, 278.7, 82.8	77.9, 101.2, 139.6	162.1, 162.1, 208.9	160.8, 160.8, 137.6
α , β , γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 120.0	90.0, 91.1, 90.0	90.0, 90.0, 120.0	90.0, 90.0, 90.0
Resolution (Å)	87.97-2.46 (2.52-2.46)	80.44-2.80 (2.88-2.80)	81.97-2.60 (2.66-2.60)	83.73-3.80 (4.01-3.80)	87.66-2.80 (2.90-2.80)
<i>R</i> _{merge}	0.097 (1.521)	0.072 (1.13)	0.173 (1.64)	0.110 (2.13)	0.127 (1.88)
<i>I</i> / σ <i>I</i>	12.9 (1.4)	10.2 (1.5)	5.5 (1.4)	10.8 (1.3)	10.3 (0.9)
Completeness (%)	100 (100)	100 (100)	100 (99.9)	100 (100)	100 (100)
Redundancy	9.7 (8.4)	3.9 (3.9)	3.8 (3.8)	11.2 (11.3)	14.7 (14.9)
Refinement					
Resolution (Å)	63.98-2.46 (2.50-2.46)	46.05-2.80 (2.84-2.80)	47.48-2.55 (2.58-2.55)	48.45-3.77 (3.88-3.77)	49.98-2.81 (2.87-2.81)
No. reflections	66619	59106	65091	32124	61365
<i>R</i> _{work} / <i>R</i> _{free}	0.205/0.240	0.237/0.268	0.241/0.289	0.294/0.317	0.256/0.224
No. atoms					
Protein	4622	10862	12144	11348	4640
Ligand/ion	16	0	56	0	119
Water	94	71	25	0	0
B-factors					
Protein	65.23	106.55	83.12	276.48	100.26
Ligand/ion	71.89	N/A	89.31	N/A	106.48
Water	60.59	65.00	67.10	N/A	N/A
R.m.s deviations					
Bond lengths (Å)	0.009	0.006	0.005	0.005	0.003
Bond angles (°)	1.063	0.828	1.109	1.024	1.057

*Highest resolution shell is shown in parentheses.

†One crystal was used for each structure.

CAREERS

NON-PROFIT WORK Films, science guidelines and emergency medicine all in a day **p.245**

SELF-ADVOCACY Five tips to get your paper noticed go.nature.com/g6zjxc

NATUREJOBS For the latest career listings and advice www.naturejobs.com

KYLE LARSON



Kyle Larson and his team trek over a mountain pass in Nepal, one of many adventurous expeditions that have allowed him to take his research outdoors.

FIELDWORK

Extreme research

Cavers, divers and climbers take their science to strange and wonderful places.

BY EMILY SOHN

Marina Elliott never planned to apply her outdoor adventure skills to a career in research and exploration. But in October 2013, she saw an advertisement for a project in South Africa that called for cave explorers with archaeological experience who were also small enough to squeeze through a narrow passageway to excavate an underground chamber. She was startled by how perfectly qualified she was for the job.

Already an avid rock climber and spelunker,

Elliott was then finishing up a PhD in biological anthropology at Simon Fraser University in Vancouver, Canada. She had worked on excavations in remote places, including Siberia and northern Alaska. And she had the flexibility to drop everything to spend a month in Africa.

She joined a team of five other women. One by one, they shimmied through a 12-metre-long chute with an 18-centimetre-wide pinch point, and they emerged with more than 1,500 fossils from 15 skeletons of a previously unknown species of ancient hominin called *Homo naledi*. The discovery helped her to land her current

postdoc position in biological anthropology at the University of the Witwatersrand in Johannesburg, and she now leads a team of six cavers who continue to explore the region.

Although Elliott's path required a dose of serendipity, her experience illustrates one of the many ways that scientists can combine a love for outdoor adventure with their career. Researchers who pursue extreme fieldwork say that the discoveries they make along the way provide a lifetime of adventure tales and shape their careers in positive ways.

Combining an extracurricular passion ►

► for the outdoors with a high-stakes career, however, also brings complications, including the risk that nothing will go as planned. When preparing for fieldwork, ‘adventure’ researchers need to be particularly careful with logistics to ensure success — and their own survival. In addition, they often need to acquire specialized insurance and build a safety net of teammates and strategies to deal with inevitable obstacles. Even with the best-laid plans, disaster can still come in many forms — from violent weather and political strife to crippling injuries. Flexibility and quick thinking can be the difference between a productive trip and a waste of time — or worse.

RISK MANAGEMENT

Stacy Kim, a marine ecologist at Moss Landing Marine Laboratories in California, regularly dives beneath the ice in Antarctica to study how human pollution affects life on the sea floor. On one trip, she spun out on a snowmobile and dislocated her shoulder. The weather was bad, and the remoteness of the location meant that it took several days for medical personnel to reach her by helicopter. For the rest of the trip, she was stuck on top of the ice, doing lab work instead of going underwater. Her teammates did the diving instead. “You try to make sure no single person is completely irreplaceable,” she says.

When injuries and other obstacles occur, they can prematurely end expeditions — but stopping early is not always an option when a research agenda is involved. So, like Kim, many scientists who work in extreme locations try to factor in more time, gear and logistical support than they would for trips done purely for fun.

Ecologist Catherine Cardelús of Colgate University in Hamilton, New York, does most of her field work in the tree canopies, a task that requires climbing up ropes while battling jungle heat and fending off biting insects. On each climb, she lugs a heavy pack filled with sample-collecting tags and bags, tape measures, notebooks, walkie-talkies, water, lunch and other supplies for days of work that can keep her in the trees for up to seven hours at a time.

Cardelús says that her field season lasts two or three times longer than those of scientists whose research occurs on the ground. Some days of work are inevitably lost because of rain or wind, so she always makes sure to hit the trees immediately on good days. But because it is so exhausting, nobody is allowed to do work in the canopy for more than three days in a row. “You have to be incredibly flexible and forgiving,” says Cardelús. “You wake up in the morning and there’s a shut-out rain, so you have to say, ‘Oh well, let’s punch in some data. Let’s do lab work.’ You constantly have to be open to the possibility that you can’t do what you need to do.”

It can help to consider unplanned diversions as opportunities, says Kyle Larson, a structural geologist at the University of British Columbia’s Okanagan campus in Kelowna. In 2014 on a trip to Nepal, Larson and his team were stopped by snow at an elevation of 3,000 metres — only halfway to their planned destination.

“You constantly have to be open to the possibility that you can’t do what you need to do.”

They collected what samples they could, but they were unable to gather data for a project of a student on the trek. She had to abandon the Nepal study and later did fieldwork in Saskatchewan instead, which led to some significant findings. “If you end up going to places you didn’t expect to go,” Larson says, “you can discover things you didn’t expect to discover.”

Safety is also a major concern when working on the edge, and adventurous scientists recommend erring on the side of caution, both to protect team members and to sustain funding. After all, it can take extra effort to persuade a granting agency to give money to support a dangerous, lengthy or team-heavy expedition that may not go as intended.

Early in his career, ecologist Douglas Larson (no relation to Kyle), now an emeritus professor at the University of Guelph, Canada, included climbing ropes in a purchase order so that he could rappel down cliffs in the Niagara Escarpment and study what turned out to be extremely ancient trees living there — a discovery so surprising that forest ecologists in the region disparaged the results before they saw samples.

The university’s director of safety and security told Larson that even one accident would shut his project down. From then on, he committed to extra precautions and redundancies in his equipment, so that even if someone had a heart attack, he or she could be pulled up. Instead of having two or three points of webbing attached to the anchors at a time, as do most recreational climbers, he and his group use four or five.

EXTREME BACK-UP

Strong safety nets become essential in this line of work. Elliott’s dig in South Africa had a medic on standby 24 hours a day. They notified several groups about their plans, including the South African military and a mining-rescue organization, so that they could get help quickly in case of a confined-space emergency. Many adventure-researchers buy specialized insurance that can provide rescue help in remote locations or that can cover accidents related to their work.

But no insurance company can charter a helicopter-rescue mission to a place such as K2, the second-tallest mountain in the world, where geologist Mike Searle of the UK’s University of Oxford has conducted research. He recommends building relationships with the locals in remote regions, and he always gets to know his porters. “If you’re in trouble, those are the guys who are going to carry you down.”

It also helps to be in shape, which is business as usual for many researchers who scale glaciers, climb mountains or dive to the ocean floor in frigid water. To prevent injury and to stay flexible and strong for caving, Elliott runs, hikes and takes exercise classes that combine ballet, Pilates and gymnastics. Kyle Larson lifts weights six days a week when he is not trekking. Kim freedives for fun, often down to

JOYS OF DISCOVERY

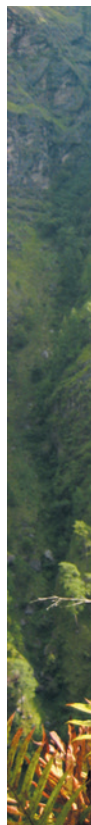
Views from the other side

Ecologist Douglas Larson, emeritus professor at the University of Guelph in Ontario, had just begun to belay a colleague over a cliff in Ontario when some weasels emerged from a nearby hole. As the colleague descended out of sight, a mother ermine and her 13 babies ran over Larson. “All I could do was scream out, ‘I’m being swarmed by ermine!’” he says. “It was one of those exquisite experiences that happens and disappears again.”

Despite its dangers and frustrations, extreme fieldwork can produce incredible, once-in-a-lifetime memories for scientists who put themselves out there. Kyle Larson (no relation to Douglas), a structural geologist at the University of British Columbia in Kelowna, once brought his father along as a field assistant on an expedition to the Himalayas. They hiked ahead of their porters, and after topping a 5,000-metre-tall peak, they stopped in a small settlement. Snow was

falling, and they were getting chilled. Then an elderly Tibetan woman poked her head out of a house and beckoned them inside. She gave them salty yak-butter tea. And although they could not communicate much, they sat around the kitchen’s cooking fire for the next hour, sharing the warmth until the rest of the team appeared. “Those things happen on every trip,” Larson says. “From a cultural perspective, it becomes very enlightening.”

Best of all, say adventurous researchers, are the rewards that come with pushing the limits of discovery. Dangling from cliff faces in Canada and France, Doug Larson has discovered tiny, 1,000-year old trees. Up in the treetops, ecologist Catherine Cardelús, an associate professor at Colgate University in Hamilton, New York, always marvels at the perspective that she gains when she doesn’t have to keep her eyes glued on the trail for snakes. “Oh my gosh,” she says. “Every day I see things most people don’t see.” **E.S.**





Kyle Larson on a research trip to Nepal.

LAURENT GODIN

18 metres or so. And Searle, now aged 61, bikes to work every day. He also climbs, swims and surfs. “You can’t climb mountains,” he says, “if you’re a couch potato.”

One summer day in 1998, Cardelús was dangling from a rope some 24 metres above ground, near the top of a tree in the Costa Rican jungle, when two howler monkeys began to make aggressive motions. Crouched about three metres away, they were shaking branches and baring their teeth with arms outstretched, ready to leap. “I thought, ‘Oh my God, here they come,’” she says. Then she heard “an ancestral guttural sound” — not from the two monkeys, but from her husband, who was working nearby. The monkeys scattered.

Cardelús — who no longer climbs when monkeys are nearby — has experienced many such close calls that include run-ins with snakes, ants, bees and tarantulas (see ‘Views from the other side’). “Each time you climb a tree for the first time, you have to be prepared to evacuate within 15 seconds,” she says. “It’s always exciting getting into the canopy. And it’s just as euphoric to get to the ground.”

Wildlife is not the only source of heart-thumping adventure. One afternoon in the spring of 2011, Kyle Larson crested a mountain pass in Nepal to discover a steep, nearly sheer descent buried in waist-deep snow. The team could see neither the trail nor what they were stepping on. Last year, he arrived in the country’s Makalu region on the heels of a

busy storm season that had dumped metres of snow on the region. Piles of snow reached the rooftops, and trekking was treacherous. “Trying to walk down through that was scary,” he says. “There were lots of bruised knees and falls.”

A certain level of psychological preparation is crucial for working in extreme environments that are, by nature, full of surprises. And that process often starts before the expedition begins. For Elliott, the idea of squeezing through an extremely tiny space presented the first mental obstacle. When Lee Berger, the palaeoanthropologist who recruited Elliott and her fellow cavers for the South Africa excavation, told them that they would need to fit through a small gap, “all of us ran around our houses measuring furniture and stuffing ourselves under it”, she says. She could wedge herself into the space just by expanding her lungs.

As she applied, Elliott worried that she wasn’t qualified enough or that she had screwed up the Skype interview. She continued to doubt herself even after arriving on site. On the first day of reconnaissance, she looked into the 12-metre-long vertical chute that the team was to descend. If someone were to get hurt, medics would have to tend to the injured person until she healed enough to get out on her own. “Psychologically, that was quite trying,” she says. “I remember looking down this shark’s maw of rock, and you can’t see where you’re going because it’s not a straight line, and thinking, ‘Oh, gee, perhaps I’ve miscalculated my own skill set.’”

Because academic courses typically do not cover the ins and outs of survival, Elliott relied instead on years of hands-on experience and prior training that had taught her to remain calm enough to deal with unexpected circumstances. Long before she took on the caving job in South Africa, she had worked as a field guide for an adventure-tour company in the Rocky Mountains and earned a certificate in wilderness first aid. Both equipped her with survival and decision-making skills.

It is impossible to predict every emergency, she says. But one can learn to think quickly and clearly in any situation. “What you can prepare for,” she says, “is the mental stability to say, ‘OK, what do we need to do next? Who needs to do it?’”

Elliott also advises young researchers to pursue all of their life passions, even if they seem completely unrelated. She started out studying veterinary medicine before earning a PhD in anthropology and landing the career-changing excavation post in South Africa. “My take-home message is, don’t fuss if your career or life path appears to be a little bit circuitous,” she says. “You never really know where any given skill set or experience might lead you.” ■

Emily Sohn is a freelance journalist in Minneapolis, Minnesota.

TRADE TALK

Science integrator



Lana Gent is a director of science at the American Heart Association in Dallas, Texas, where she coordinates networks of volunteers, the drafting of science guidelines for emergency medicine

and the production of instructional videos on first aid.

What do you do?

I help to gather input from resuscitation scientists around the world who evaluate the scientific evidence that goes into the creation of our resuscitation guidelines. I’m not usually the one who is giving the talk or is the first author, but it is my team that ensures that those experts are able to do the presentation or create the publication.

How did you learn about this job?

Our lab was in a crunch for money and that made me think about what I wanted to do. Do I continue on this pathway of being a traditional lab scientist? Colleagues were encouraging me to be a medical-science liaison — a professional who teaches physicians how new medicines work — and I was going to interviews. During that process, I was contacted by a recruiter representing the American Heart Association. I didn’t think I was the best fit; I didn’t have a resuscitation-science background or management experience, but the recruiters knew that I had transferable skills. I convinced everyone that if I could learn stereotactic brain surgery in mice, I could learn resuscitation science.

Why is this job right for you?

No one goes to school for the type of position I have. You wear a lot of hats. The hiring manager could see that I took initiative and was passionate. I had shown that I could take on new challenges and bring people together.

What has the job taught you?

I had to learn to be resilient and inquisitive and not walk along just one path. Sometimes as scientists we pride ourselves on being contemplative, and the greatest skill set here is to simplify information and to be quick on your feet. ■

INTERVIEW BY MONYA BAKER

This interview has been edited for length and clarity. For more, see go.nature.com/iyssud.

PROJECT DAFFODIL

Hothouse flowers.

BY SYLVIA SPRUCK WRIGLEY

The Daffodil Project is a joint venture between the United Nations, ESA and Friends of the Elderly. I am their 17th civilian recruit. For obvious reasons, the entire project is extremely hush-hush.

My daughter hates the Project.

"Maman," she said. Amélie visits once a week. "Maman, let me take you out of here. There's no mission." I thought we'd had our funding cut. Then I saw the look her son Daniel gave her. Poor Amélie. She doesn't think I should be considering space travel, not at my age. Every visit she's tried to tell me that the Daffodil Project doesn't exist. We've had this conversation so many times.

"What's the harm," Daniel whispered, like I can't hear. At least he understands I'm not abandoning them, or if I am, well... funny, that my grandson would be the one to understand.

"Tell me again, grand-mère, how are they going to get all of you to Mars? Will you get a window seat? Can you send me pictures?"

Amélie scowled. She's always hid her fear under a face full of anger. She can't bear the thought of my leaving her, even though she knows it will happen someday. I wish I could comfort her like I did when she was a little girl scared of thunder. She'll be 50 in February but she's still my Mélie.

"I don't think the space-transport will have windows," I told Daniel. "It's not a pleasure cruise. But once we arrive, we'll be able to see Mars and get instructions from Earth. We have some sort of data transfer over satellite for systems monitoring. I bet we'll even have voice data. I'll have to ask Jacques..."

The nurse walking past smiled. "Aren't you the clever one, Madame?"

That was when Amélie stormed out, threatening to shut the place down as she dragged her son out the door. I didn't mind. I could see Claudine setting up the card table for a game of belote.

These are my friends now. Paul spread chocolates on the card table, a gift from his granddaughter. Awful American chocolate, but we all took a piece anyway. We'll be on hard rations soon



enough. We're trying to emulate life on the space station so the food is not very good, all government approved. The thing is, when we get up there, there's no coming back. So if an addiction to croissants or cognac is going to break your heart, you'd better find out now.

Or if leaving for Mars is going to break your daughter's heart.

Claudine was dealer, as always. She's 83, the eldest. She's also the best belote player, although Jacques swears she cheats.

There's two dozen of us here in the home, if I don't count Monsieur Coulon, which I don't because he's really just a bundle of robes and drool. I guess they keep him with us as a test, to see how we cope when someone loses it completely but doesn't die. The nurses said not everyone can go to space. So maybe they're sending a dozen, or less?

We're all very well behaved because we want to get chosen. It's true, Mars's surface is too cold for space colonies. But heating up a planet is not a problem; it just takes time. It's on the television every evening how we've overheated this one, after all. Jacques and I have worked out the super-secret plan: send a small team to Mars to oversee operations. Repeat as necessary until the planet is inhabitable and then invite the masses. Voilà: new homeland.

Amélie thinks I'm deluded. She half expects me to transfer my life's savings to a con-man or something. But honestly, my life's savings are a €800-per-month pension and the house

that Antoine and I bought in 1972. Really, what do I have to lose?

What Amélie doesn't understand is that we aren't colonizing. We're Project Daffodil, the first inhabitants of chilly Mars. Jacques and I, we've worked it all out. They can't afford to keep ferrying people back and forth. But we've got some time left, five years, maybe ten. Our job is to keep an eye on things and wait to die.

If you want to save money with a one-way trip, it makes perfect sense to recruit from a nursing home, now doesn't it?

This is why we practise survival skills. We can't live on the surface, obviously, so we'll have a dome. It's going to be small and uncomfortable, like our shared rooms here: just enough space for two beds and a cupboard, no chance of privacy.

If it is just four of us, I hope I get to go with Claudine and Paul and Jacques. We play belote every afternoon and tell stories and have a grand time when we can convince one of the grandchildren to bring us pastis. I guess they won't give us pastis up there, but we'll take the cards. We'll get by.

Amélie keeps making up excuses to try to get me out of here. She's frantic for me to come and live with her now, leave the home and the Project.

I told her that I'll think about it, but Jacques thinks they are confirming the final crew soon. Maybe even tonight.

Since Antoine died, every morning I wake up and remember that I'm decrepit. That he abandoned me to live my old age alone. Every morning I remember that there's no future, my entire life is in the past. So, what do I do?

I could deaden it with pastis until someone carts me away.

Or I could go to Mars.

Tell Amélie that I will miss her. ■

Sylvia Spruck Wrigley was born in Germany and spent her childhood in Los Angeles. She now splits her time between South Wales and Andalucia. Her fiction was nominated for a Nebula in 2014 and her short stories have been translated into more than a dozen languages. Her first novella, Domnall and the Borrowed Child, is out now as part of the Tor.com novella imprint. You can find out more about her at www.intrigue.co.uk.

ILLUSTRATION BY JACEY